TRANSGENIC MICE CONTAINING THYROID STIMULATING HORMONE RECEPTOR (TSH-R) GENE DISRUPTIONS

Related Applications

5   This application claims priority to U.S. Provisional Application No. 60/243,895, filed October 26, 2000, the entire contents of which are incorporated herein by reference.

Field of the Invention

The present invention relates to transgenic animals, compositions and methods
10   relating to the characterization of gene function.

Background of the Invention

Many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second
15   messengers such as cAMP. The membrane protein gene superfamily of G-protein coupled receptors (GPCRs) include a wide range of biologically active receptors, such as hormone, viral, growth factor and neuroreceptors. GPCRs have been characterized as having seven putative transmembrane domains (designated TM1, TM2, TM3, TM4, TM5, TM6, and TM7), which are believed to represent transmembrane α-helices
20   connected by extracellular or cytoplasmic loops. Most G-protein coupled receptors have single conserved cysteine residues in each of the first two extracellular loops that form disulfide bonds that are believed to stabilize functional protein structure. G-protein coupled receptors can be intracellularly coupled by heterotrimeric G-proteins to various intracellular enzymes, ion channels and transporters. Different G-protein α-subunits
25   preferentially stimulate particular effectors to modulate various biological functions in a cell.

Over the past 15 years, nearly 350 therapeutic agents targeting 7-transmembrane receptors have been successfully introduced onto the market. As these receptors have an established history as therapeutic targets, a clear need exists for identification and
30   characterization of GPCRs which can play a role in preventing, ameliorating or correcting dysfunctions or diseases.

One important subfamily of the GPCRs is the thyroid stimulating hormone receptor (TSH-R) GPCRs. Thyroid-stimulating hormone (TSH), also known as thyrotropin, is secreted from cells in the anterior pituitary called thyrotrophs, finds its receptors on epithelial cells in the thyroid gland, and stimulates that gland to synthesize

5 and release thyroid hormones. TSH is a glycoprotein hormone (MW = 28 kD) composed of two subunits that are non-covalently bound to one another. The alpha subunit of TSH is also present in two other pituitary glycoprotein hormones, follicle-stimulating hormone (FSH) and luteinizing hormone (LH). Each of these hormones also has a unique beta subunit, which provides receptor specificity. In other words, TSH is composed of an

10 alpha subunit bound to a TSH beta subunit, and TSH associates only with its own receptor.

The most important controller of TSH secretion is thyroid-releasing hormone (TRH). Thyroid-releasing hormone is secreted by hypothalamic neurons into hypothalamic-hypophyseal portal blood, finds its receptors on thyrotrophs in the anterior

15 pituitary and stimulates secretion of TSH. One interesting aspect of thyroid-releasing hormone is that it is only three amino acids long. Its basic sequence is glutamic acid-histidine-proline, although both ends of the peptide are modified. Secretion of thyroid-releasing hormone, and hence, TSH, is inhibited by high blood levels of thyroid hormones in a classical negative feedback loop.

20 Thyroid function is controlled by TSH secreted by the anterior pituitary. When pituitary TSH binds to the receptors on the cell surface of the thyroid epithelial cells it activates adenylate cyclase, and the resultant increase in intracellular cyclic AMP brings about most of the effects of TSH including the stimulation of thyroid hormone synthesis and the regulation of thyroid growth and development. The secretion of TSH is in turn

25 regulated by a direct inhibitory feedback of high levels of circulating thyroid hormones. A disturbance in this regulation takes place in Graves' disease, which is characterized by hyperthyroidism, goiter and exophthalmos.

The TSH-R gene was first cloned and sequenced by Nagayama et al. (*Biochem. Biophys. Res. Commun.* 165: 1184-1190 (1989)) and by Libert et al. (*Biochem. Biophys.*

30 *Res. Commun.* 165: 1250-1255 (1989)). The TSH receptor differs from the luteinizing hormone/choriogonadotropin receptor (LHCGR) by the presence of 2 unique insertions

2

of 8 and 50 amino acids in the extracellular domain. Wadsworth et al. (*Science* 249: 1423-1425 (1990)) showed that the 8-amino acid tract near the amino terminus of the TSH receptor is an important site of interaction with both TSH and autoantibodies against the TSH receptor (thyroid-stimulating immunoglobulins, TSI). Either deletion or

5   substitution of this region abolished the interaction, whereas a deletion of the 50-amino acid tract had no effect. Vassart et al. (*Trends Endocr. Metab.* 2: 151-156 (1991)) reviewed the molecular genetics of TSH-R. The high sequence homology with the LHCG receptor, which is composed of a single polypeptide chain, led many to suppose a similar structure for the TSH receptor. However, Loosfelt et al. (*Proc. Nat. Acad. Sci.*

10   89: 3765-3769 (1992)) presented evidence for a heterodimeric structure of TSH-R. The extracellular (hormone-binding) alpha subunit had an apparent molecular mass of 53 kD, whereas the membrane-spanning beta subunit seemed heterogeneous and had an apparent molecular mass of 33 to 42 kD. Human thyroid membranes have been found to contain 2.5 to 3 times as many beta subunits as alpha subunits; however, the two subunits

15   probably derive from a single gene since a single reading frame was demonstrated by cDNA cloning and sequencing.

Graves et al. (*J. Clin. Endocr. Metab.* 84: 2177-2181 (1999)) used epitope-mapped monoclonal and polyclonal antibodies to TSH-R as immunoblot probes to detect and characterize the molecular species of the receptor present in normal human thyroid

20   tissue. In reduced membrane fractions, both full-length (uncleaved) holoreceptor and cleavage-derived subunits of the holoreceptor were detected. Uncleaved holoreceptor species included a nonglycosylated form of apparent molecular mass 85 kD and two glycosylated forms of approximately 110 and 120 kD. The membranes also contained several forms of cleavage-derived TSH-R alpha and beta subunits. Alpha subunits were

25   detected by antibodies to epitopes localized within the N-terminal end of the TSH-R ectodomain and migrated diffusely between 45 and 55 kD, reflecting a differentially glycosylated status. Several species of beta subunit were present, the most abundant having apparent molecular masses of 50, 40, and 30 kD. The authors concluded that posttranslational processing of the TSH-R occurs in human thyroid tissue and involves

30   multiple cleavage sites.

McBride et al. (*Am. J. Hum. Genet.* 41: A177 (1987)) mapped what was believed to be the structural gene for TSH receptor using cDNA clones isolated from human thyroid carcinoma by means of immunoscreening with antiserum from a patient with Graves disease. They reported that a 1.8-kb full-length human TSH receptor cDNA was obtained by screening a thyroid gt11 library. The complete DNA sequence of this insert had a 1.4-kb open reading frame encoding a polypeptide with hallmarks of a membrane-spanning protein. Akamizu et al. (*Proc. Nat. Acad. Sci.* 87: 5677-5681 (1990)) mapped the TSH-R gene to human chromosome 14 by study of somatic cell hybrid DNAs and to mouse chromosome 12 by linkage studies in interspecies backcross mice. By *in situ* hybridization, Rousseau-Merck et al. (*Genomics* 8: 233-236 (1990)) and Libert et al. (*Cytogenet. Cell Genet.* 54: 82-83 (1990)) regionalized the gene to 14q31. The homologous gene in the mouse was confirmed to be located on chromosome 12 (Wilkie et al., *Genomics* 18: 175-184 (1993)).

By PCR amplification of specific cDNA, Feliciello et al. (*Lancet* 342: 337-338 (1993)) demonstrated that mature TSH receptor mRNA is expressed in the retroorbital tissue of both healthy subjects and patients with Graves disease. Of other tissues and cells tested, only thyroid tissue expressed the TSH-R mRNA. Bahn et al. (*J. Endocr. Invest.* 16 (suppl. 2-6): 30 (1993)) found a point mutation in TSH receptor mRNA prepared from retrobulbar fibroblasts in one of seven patients with severe Graves ophthalmopathy. The mutation resulted in a single amino acid substitution in a putative immunogenic region of the TSH receptor protein (Murakami and Mori, *Biochem. Biophys. Res. Commun.* 171: 512-518 (1990)). Heldin et al. (*J. Clin. Endocr. Metab.* 73: 1374-1376 (1991)) described other point mutations in TSH receptor mRNA in the thyroid of a patient with Graves disease. Rodien et al. (*New Eng. J. Med.* 339: 1823-1826 (1998)) described a gain-of-function mutation of the TSH-R gene as the cause of familial gestational hyperthyroidism. The mutation rendered the thyrotropin receptor hypersensitive to chorionic gonadotropin.

Lazar et al. (*J. Clin. Endocr. Metab.* 84: 3228-3234 (1999)) studied the expression of TSH-R in ninety human thyroid tissues. mRNAs were extracted from 43 thyroid carcinomas (38 papillary and 5 follicular), 24 cold adenomas, 5 Graves thyroid tissues, 8 toxic adenomas, and 5 hyperplastic thyroid tissues; 5 normal thyroid tissues were used as

4

reference. A kinetic quantitative PCR method, based on the fluorescent TaqMan methodology and real-time measurement of fluorescence, was used. TSH-R expression was normal in most tissues studied and was decreased in only some thyroid carcinomas.

The complete mRNA cds for the murine (Mus musculus BALB/cBY) TSH-R gene has been deposited in GenBank and bears Accession number U02602 and GI/NID number 575923. The TSH-R gene sequence is described in S.A. Stein et al., *Mol. Endocrinol.* 8(2): 129-38 (1994).

The hyt/hyt hypothyroid mouse has an autosomal recessive, fetal-onset, severe hypothyroidism related to TSH hyporesponsiveness and associated with elevated TSH (S.A. Stein et al., *Mol. Endocrinol.* 8(2): 129-38 (1994)). It had been previously suggested that the hypothyroidism and TSH hyporesponsiveness may result from a mutation in the hyt/hyt TSH receptor of the thyroid gland. Based on DNA sequencing of the entire coding region of the TSH-R gene from the wild-type BALB/cBY +/+ mouse, the +/+ TSH-R is 92% and 94% identical at the nucleotide and amino acid residue levels, respectively, compared to the rat TSH-R gene. The coding region of the hyt/hyt TSH-R, compared to that of the +/+ TSH-R, has a single base change (C is replaced by T in the hyt/hyt mutant) at nucleotide position 1740.

Given the importance of GPCRs, such as TSH-R, a clear need exists for their identification and characterization, which can play a role in preventing, ameliorating or correcting dysfunctions or diseases.

Summary of the Invention

The present invention generally relates to transgenic animals, as well as to compositions and methods relating to the characterization of gene function.

The present invention provides transgenic cells comprising a disruption in a TSH-R gene. The transgenic cells of the present invention are comprised of any cells capable of undergoing homologous recombination. Preferably, the cells of the present invention are stem cells and more preferably, embryonic stem (ES) cells, and most preferably, murine ES cells. According to one embodiment, the transgenic cells are produced by introducing a targeting construct into a stem cell to produce a homologous recombinant, resulting in a mutation of the TSH-R gene. In another embodiment, the transgenic cells are derived from the transgenic animals described below. The cells derived from the

transgenic animals includes cells that are isolated or present in a tissue or organ, and any cell lines or any progeny thereof.

The present invention also provides a targeting construct and methods of producing the targeting construct that when introduced into stem cells produces a homologous recombinant. In one embodiment, the targeting construct of the present invention comprises first and second polynucleotide sequences that are homologous to the TSH-R gene. The targeting construct may also comprise a polynucleotide sequence that encodes a selectable marker that is preferably positioned between the two different homologous polynucleotide sequences in the construct. The targeting construct may also comprise other regulatory elements that can enhance homologous recombination.

The present invention further provides non-human transgenic animals and methods of producing such non-human transgenic animals comprising a disruption in a TSH-R gene. The transgenic animals of the present invention include transgenic animals that are heterozygous and homozygous for a null mutation in the TSH-R gene. In one aspect, the transgenic animals of the present invention are defective in the function of the TSH-R gene. In another aspect, the transgenic animals of the present invention comprise a phenotype associated with having a mutation in a TSH-R gene. Preferably, the transgenic animals are rodents and, most preferably, are mice. In accordance with the present invention, transgenic mice having a disruption in the TSH-R gene exhibit at least one of the following phenotypes (relative to wild-type controls): reduced body size and/or abnormal shape; "dwarfism" (i.e., homozygotes smaller than their wild-type counterparts); a growth disorder; small head, eyes, and/or ears; abnormal and/or short or small snouts; and abnormal (e.g., hunched) posture. The transgenic animals of the present invention may alternatively or additionally exhibit at least one of the following phenotypes (relative to wild-type controls): abnormal activity level (e.g., hypo- or hyper-activity); abnormal metabolic rate; and reduced responsiveness or non-responsiveness. The transgenic animals of the present invention may alternatively or additionally exhibit at least one of the following phenotypes (relative to wild-type controls): low body weights; short body lengths; low organ (e.g., spleen, liver, kidneys, heart, and thymus gland) weights; and low organ weight:body weight ratios, especially for the spleen, liver, and kidneys. The transgenic animals of the present invention may alternatively or

6

additionally exhibit at least one of the following phenotypes (relative to wild-type controls): small thymus glands; small skeletal muscle; reduced subcutaneous body fat; small or absent seminal vesicles; and malformed femurs. The transgenic animals of the present invention may alternatively or additionally exhibit at least one of the following phenotypes (relative to wild-type controls): histopathological lesions on organs; small thyroid glands (especially with small follicle size); adenohypophysis of the pituitary gland; large and vacuolated cells (e.g., acidophils) in the anterior pituitary gland; dysplasia of the epiphyses of the femur, tibia and/or stifle joint (especially where characterized by reduced and/or patchy ossification); reduced cellularity in bone marrow; hypoplastic thymus glands; immature kidneys, especially with small glomeruli; immature testes, especially with hypospermatogenesis; hypoplastic interstitial (Leydig) cells; and oligospermia in the epididymides. The transgenic animals of the present invention may alternatively or additionally exhibit at least one of the following phenotypes (relative to wild-type controls): elevated blood urea nitrogen levels; and infertility. Immaturity or smallness of certain organs may be indicative of arrested development and/or reduced developmental rate.

The present invention also provides methods of identifying agents capable of affecting a phenotype of a transgenic animal. For example, a putative agent is administered to the transgenic animal and a response of the transgenic animal to the putative agent is measured and compared to the response of a "normal" or wild-type mouse, or alternatively compared to a transgenic animal control (without agent administration). The invention further provides agents identified according to such methods. The present invention also provides methods of identifying agents useful as therapeutic agents for treating conditions associated with a disruption or other mutation (including naturally occurring mutations) of the TSH-R gene.

The present invention further provides a method of identifying agents having an effect on TSH-R expression or function. The method includes administering an effective amount of the agent to a transgenic animal, preferably a mouse. The method includes measuring a response of the transgenic animal, for example, to the agent, and comparing the response of the transgenic animal to a control animal, which may be, for example, a wild-type animal or alternatively, a transgenic animal control. Compounds that may have

7

an effect on TSH-R expression or function may also be screened against cells in cell-based assays, for example, to identify such compounds.

The invention also provides cell lines comprising nucleic acid sequences of a TSH-R gene. Such cell lines may be capable of expressing such sequences by virtue of operable linkage to a promoter functional in the cell line. Preferably, expression of the TSH-R gene sequence is under the control of an inducible promoter. Also provided are methods of identifying agents that interact with the TSH-R gene, comprising the steps of contacting the TSH-R gene with an agent and detecting an agent/TSH-R gene complex. Such complexes can be detected by, for example, measuring expression of an operably linked detectable marker.

The invention further provides methods of treating diseases or conditions associated with a disruption in a TSH-R gene, and more particularly, to a disruption in the expression or function of the TSH-R gene. In a preferred embodiment, methods of the present invention involve treating diseases or conditions associated with a disruption in the TSH-R gene's expression or function, including administering to a subject in need, a therapeutic agent that affects TSH-R expression or function. In accordance with this embodiment, the method comprises administration of a therapeutically effective amount of a natural, synthetic, semi-synthetic, or recombinant TSH-R gene, TSH-R gene products or fragments thereof as well as natural, synthetic, semi-synthetic or recombinant analogs.

The present invention also provides compositions comprising or derived from ligands or other molecules on compounds that bind to or interact with the TSH receptor, including agonists or antagonists of TSH-R. Such agonists or antagonists of TSH-R include antibodies and antibody mimetics, as well as other molecules that can readily be identified by routine assays and experiments well known in the art.

The present invention further provides methods of treating diseases or conditions associated with disrupted targeted gene expression or function, wherein the methods comprise detecting and replacing through gene therapy mutated or otherwise defective or abnormal TSH-R genes.

8

<u>Definitions</u>

The term "gene" refers to (a) a gene containing at least one of the DNA sequences disclosed herein; (b) any DNA sequence that encodes the amino acid sequence encoded by the DNA sequences disclosed herein and/or; (c) any DNA sequence that hybridizes to

5   the complement of the coding sequences disclosed herein. Preferably, the term includes coding as well as noncoding regions, and preferably includes all sequences necessary for normal gene expression including promoters, enhancers and other regulatory sequences.

The terms "polynucleotide" and "nucleic acid molecule" are used interchangeably to refer to polymeric forms of nucleotides of any length. The polynucleotides may

10   contain deoxyribonucleotides, ribonucleotides and/or their analogs. Nucleotides may have any three-dimensional structure, and may perform any function, known or unknown. The term "polynucleotide" includes single-, double-stranded and triple helical molecules. "Oligonucleotide" refers to polynucleotides of between 5 and about 100 nucleotides of single- or double-stranded DNA. Oligonucleotides are also known as oligomers or oligos

15   and may be isolated from genes, or chemically synthesized by methods known in the art. A "primer" refers to an oligonucleotide, usually single-stranded, that provides a 3'-hydroxyl end for the initiation of enzyme-mediated nucleic acid synthesis. The following are non-limiting embodiments of polynucleotides: a gene or gene fragment, exons, introns, mRNA, tRNA, rRNA, ribozymes, cDNA, recombinant polynucleotides,

20   branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes and primers. A nucleic acid molecule may also comprise modified nucleic acid molecules, such as methylated nucleic acid molecules and nucleic acid molecule analogs. Analogs of purines and pyrimidines are known in the art, and include, but are not limited to, aziridinycytosine, 4-acetylcytosine,

25   5-fluorouracil, 5-bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxymethyl-aminomethyluracil, inosine, N6-isopentenyladenine, 1-methyladenine, 1-methylpseudouracil, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, pseudouracil, 5-pentylnyluracil and 2,6-diaminopurine. The use of uracil as a substitute for thymine in a

30   deoxyribonucleic acid is also considered an analogous form of pyrimidine.

A "fragment" of a polynucleotide is a polynucleotide comprised of at least 9 contiguous nucleotides, preferably at least 15 contiguous nucleotides and more preferably at least 45 nucleotides, of coding or non-coding sequences.

The term "gene targeting" refers to a type of homologous recombination that occurs when a fragment of genomic DNA is introduced into a mammalian cell and that fragment locates and recombines with endogenous homologous sequences.

The term "homologous recombination" refers to the exchange of DNA fragments between two DNA molecules or chromatids at the site of homologous nucleotide sequences.

The term "homologous" as used herein denotes a characteristic of a DNA sequence having at least about 70 percent sequence identity as compared to a reference sequence, typically at least about 85 percent sequence identity, preferably at least about 95 percent sequence identity, and more preferably about 98 percent sequence identity, and most preferably about 100 percent sequence identity as compared to a reference sequence. Homology can be determined using , for example, a "BLASTN" algorithm. It is understood that homologous sequences can accommodate insertions, deletions and substitutions in the nucleotide sequence. Thus, linear sequences of nucleotides can be essentially identical even if some of the nucleotide residues do not precisely correspond or align. The reference sequence may be a subset of a larger sequence, such as a portion of a gene or flanking sequence, or a repetitive portion of a chromosome.

The term "target gene" (alternatively referred to as "target gene sequence" or "target DNA sequence" or "target sequence") refers to any nucleic acid molecule or polynucleotide of any gene to be modified by homologous recombination. The target sequence includes an intact gene, an exon or intron, a regulatory sequence or any region between genes. The target gene comprises a portion of a particular gene or genetic locus in the individual's genomic DNA. As provided herein, the target gene of the present invention is a TSH-R gene, or a homolog or ortholog thereof. A "TSH-R gene" refers to a sequence comprising SEQ ID NO:1 or comprising the TSH-R sequence identified in GenBank as Accession No.: U02602; GI: 575923, or homologs or orthologs thereof.

"Disruption" of a TSH-R gene occurs when a fragment of genomic DNA locates and recombines with an endogenous homologous sequence. These sequence disruptions

10

or modifications may include insertions, missense, frameshift, deletion, or substitutions, or replacements of DNA sequence, or any combination thereof. Insertions include the insertion of entire genes, which may be of animal, plant, fungal, insect, prokaryotic, or viral origin. Disruption, for example, can alter or replace a promoter, enhancer, or splice site of a TSH-R gene, and can alter the normal gene product by inhibiting its production partially or completely or by enhancing the normal gene product's activity.

The term "transgenic cell" refers to a cell containing within its genome a TSH-R gene that has been disrupted, modified, altered, or replaced completely or partially by the method of gene targeting.

The term "transgenic animal" refers to an animal that contains within its genome a specific gene that has been disrupted by the method of gene targeting. Transgenic animals include both the heterozygote animal (*i.e.*, one defective allele and one wild-type allele) and the homozygous animal (*i.e.*, two defective alleles).

As used herein, the terms "selectable marker" or "positive selection marker" refer to a gene encoding a product that enables only the cells that carry the gene to survive and/or grow under certain conditions. For example, plant and animal cells that express the introduced neomycin resistance (*Neo^r*) gene are resistant to the compound G418. Cells that do not carry the *Neo^r* gene marker are killed by G418. Other positive selection markers are known to those of ordinary skill in the art.

A "host cell" includes an individual cell or cell culture that can be or has been a recipient for vector(s) or for incorporation of nucleic acid molecules and/or proteins. Host cells include progeny of a single host cell, and the progeny may not necessarily be completely identical (in morphology or in total DNA complement) to the original parent due to natural, accidental, or deliberate mutation. A host cell includes cells transfected with the constructs of the present invention.

The term "modulates" as used herein refers to the decrease, inhibition, reduction, increase or enhancement of TSH-R function, expression, activity, or alternatively a phenotype associated with a disruption in a TSH-R gene.

The term "ameliorates" refers to a decrease, reduction or elimination of a condition, disease, disorder, or phenotype, including an abnormality or symptom associated with a disruption in a TSH-R gene.

The term "abnormality" refers to any disease, disorder, condition, or phenotype in which a disruption of a TSH-R gene is implicated, including pathological conditions and behavioral observations.

Brief Description of the Drawings

Figure 1A shows the polynucleotide sequence for an orphan TSH-R gene (SEQ ID NO:1). Figure 1B shows the amino acid sequence for the TSH-R gene (SEQ ID NO:2).

Figures 2A-2B show the design of the targeting construct used to disrupt TSH-R genes. Figure 2B shows the sequences identified as SEQ ID NO:3 and SEQ ID NO:4, which were used as the targeting arms (homologous sequences) in the TSH-R targeting construct.

Figure 3 shows a table relating organ weight and organ weight:body weight ratios for the spleen, liver, kidney, thymus and heart of approximately 49-day old wild-type mice, heterozygous mice and homozygous mutant mice. Also shown are the body weight and body length for the same approximately 49-day old wild-type mice, heterozygous mice and homozygous mutant mice, and testes and epididymis weights for the male mice.

Figure 4 shows a table relating organ weight and organ weight:body weight ratios for the spleen, liver, kidney, thymus and heart of approximately 300-day old wild-type mice, heterozygous mice and homozygous mutant mice. Also shown are the body weight and body length for the same approximately 300-day old wild-type mice, heterozygous mice and homozygous mutant mice, and testes and epididymis weights for the male mice.

Detailed Description of the Invention

The invention is based, in part, on the evaluation of the expression and role of genes and gene expression products, primarily those associated with a TSH-R gene. Among other uses or applications, the invention permits the definition of disease pathways and the identification of diagnostically and therapeutically useful targets. For example, genes that are mutated or down-regulated under disease conditions may be

involved in causing or exacerbating the disease condition. Treatments directed at up-regulating the activity of such genes or treatments that involve alternate pathways, may ameliorate the disease condition.

5   Generation of Targeting Construct

The targeting construct of the present invention may be produced using standard methods known in the art. (*see, e.g.*, Sambrook, *et al.*, 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; E.N. Glover (eds.), 1985, *DNA Cloning: A Practical Approach*,

10   Volumes I and II; M.J. Gait (ed.), 1984, *Oligonucleotide Synthesis*; B.D. Hames & S.J. Higgins (eds.), 1985, Nucleic Acid Hybridization; B.D. Hames & S.J. Higgins (eds.), 1984, *Transcription and Translation*; R.I. Freshney (ed.), 1986, Animal Cell Culture; Immobilized Cells and Enzymes, IRL Press, 1986; B. Perbal, 1984, A Practical Guide To Molecular Cloning; F.M. Ausubel *et al.*, 1994, *Current Protocols in Molecular Biology*,

15   John Wiley & Sons, Inc.). For example, the targeting construct may be prepared in accordance with conventional ways, where sequences may be synthesized, isolated from natural sources, manipulated, cloned, ligated, subjected to *in vitro* mutagenesis, primer repair, or the like. At various stages, the joined sequences may be cloned, and analyzed by restriction analysis, sequencing, or the like.

20   The targeting DNA can be constructed using techniques well known in the art. For example, the targeting DNA may be produced by chemical synthesis of oligonucleotides, nick-translation of a double-stranded DNA template, polymerase chain-reaction amplification of a sequence (or ligase chain reaction amplification), purification of prokaryotic or target cloning vectors harboring a sequence of interest (*e.g.*, a cloned

25   cDNA or genomic DNA, synthetic DNA or from any of the aforementioned combination) such as plasmids, phagemids, YACs, cosmids, bacteriophage DNA, other viral DNA or replication intermediates, or purified restriction fragments thereof, as well as other sources of single and double-stranded polynucleotides having a desired nucleotide sequence. Moreover, the length of homology may be selected using known methods in

30   the art. For example, selection may be based on the sequence composition and complexity of the predetermined endogenous target DNA sequence(s).

The targeting construct of the present invention typically comprises a first sequence homologous to a portion or region of the TSH-R gene and a second sequence homologous to a second portion or region of the TSH-R gene. The targeting construct further comprises a positive selection marker, which is preferably positioned in between the first and the second DNA sequence that are homologous to a portion or region of the target DNA sequence. The positive selection marker may be operatively linked to a promoter and a polyadenylation signal.

Other regulatory sequences known in the art may be incorporated into the targeting construct to disrupt or control expression of a particular gene in a specific cell type. In addition, the targeting construct may also include a sequence coding for a screening marker, for example, green fluorescent protein (GFP), or another modified fluorescent protein.

Although the size of the homologous sequence is not critical and can range from as few as 50 base pairs to as many as 100 kb, preferably each fragment is greater than about 1 kb in length, more preferably between about 1 and about 10 kb, and even more preferably between about 1 and about 5 kb. One of skill in the art will recognize that although larger fragments may increase the number of homologous recombination events in ES cells, larger fragments will also be more difficult to clone.

In a preferred embodiment of the present invention, the targeting construct is prepared directly from a plasmid genomic library using the methods described in pending U.S. Patent Application Ser. No.: 08/971,310, filed November 17, 1997, the disclosure of which is incorporated herein in its entirety. Generally, a sequence of interest is identified and isolated from a plasmid library in a single step using, for example, long-range PCR. Following isolation of this sequence, a second polynucleotide that will disrupt the target sequence can be readily inserted between two regions encoding the sequence of interest. In accordance with this aspect, the construct is generated in two steps by (1) amplifying (for example, using long-range PCR) sequences homologous to the target sequence, and (2) inserting another polynucleotide (for example a selectable marker) into the PCR product so that it is flanked by the homologous sequences. Typically, the vector is a plasmid from a plasmid genomic library. The completed construct is also typically a circular plasmid.

14

In another embodiment, the targeting construct is designed in accordance with the regulated positive selection method described in U.S. Patent Application Ser. No. 60/232,957, filed September 15, 2000, the disclosure of which is incorporated herein in its entirety. The targeting construct is designed to include a PGK-*neo* fusion gene having two *lacO* sites,

5   positioned in the PGK promoter and an NLS-*lacI* gene comprising a lac repressor fused to sequences encoding the NLS from the SV40 T antigen.

In another embodiment, the targeting construct may contain more than one selectable maker gene, including a negative selectable marker, such as the herpes simplex virus tk (HSV-tk) gene. The negative selectable marker may be operatively linked to a

10   promoter and a polyadenylation signal. (*see, e.g.*, U.S. Patent No. 5,464,764; U.S. Patent No. 5,487,992; U.S. Patent No. 5,627,059; and U.S. Patent No. 5,631,153).

Generation of Cells and Confirmation of Homologous Recombination Events

Once an appropriate targeting construct has been prepared, the targeting construct

15   may be introduced into an appropriate host cell using any method known in the art. Various techniques may be employed in the present invention, including, for example, pronuclear microinjection; retrovirus mediated gene transfer into germ lines; gene targeting in embryonic stem cells; electroporation of embryos; sperm-mediated gene transfer; and calcium phosphate/DNA co-precipitates, microinjection of DNA into the

20   nucleus, bacterial protoplast fusion with intact cells, transfection, polycations, *e.g.*, polybrene, polyornithine, *etc.*, or the like (*see, e.g.*, U.S. Patent No. 4,873,191; Van der Putten, *et al.*, 1985, *Proc. Natl. Acad. Sci., USA* 82:6148-6152; Thompson, *et al.*, 1989, *Cell* 56:313-321; Lo, 1983, *Mol Cell. Biol.* 3:1803-1814; Lavitrano, *et al.*,1989, *Cell*, 57:717-723). Various techniques for transforming mammalian cells are known in the art.

25   (*see, e.g.*, Gordon, 1989, *Intl. Rev. Cytol.*, 115:171-229; Keown *et al.*, 1989, *Methods in Enzymology*; Keown *et al.*, 1990, *Methods and Enzymology*, Vol. 185, pp. 527-537; Mansour *et al.*, 1988, *Nature*, 336:348-352).

In a preferred aspect of the present invention, the targeting construct is introduced into host cells by electroporation. In this process, electrical impulses of high field

30   strength reversibly permeabilize biomembranes allowing the introduction of the construct. The pores created during electroporation permit the uptake of macromolecules

15

such as DNA. (*see, e.g.*, Potter, H., *et al.*, 1984, *Proc. Nat'l. Acad. Sci. U.S.A.* 81:7161-7165).

Any cell type capable of homologous recombination may be used in the practice of the present invention. Examples of such target cells include cells derived from

5 vertebrates including mammals such as humans, bovine species, ovine species, murine species, simian species, and ether eucaryotic organisms such as filamentous fungi, and higher multicellular organisms such as plants.

Preferred cell types include embryonic stem (ES) cells, which are typically obtained from pre-implantation embryos cultured *in vitro*. (*see, e.g.*, Evans, M. J., *et al.*,

10 1981, *Nature* 292:154-156; Bradley, M. O., *et al.*, 1984, *Nature* 309:255-258; Gossler *et al.*, 1986, *Proc. Natl. Acad. Sci. USA* 83:9065-9069; and Robertson, *et al.*, 1986, *Nature* 322:445-448). The ES cells are cultured and prepared for introduction of the targeting construct using methods well known to the skilled artisan. (*see, e.g.*, Robertson, E. J. ed. "Teratocarcinomas and Embryonic Stem Cells, a Practical Approach", IRL Press,

15 Washington D.C., 1987; Bradley *et al.*, 1986, *Current Topics in Devel. Biol.* 20:357-371; by Hogan *et al.*, in "Manipulating the Mouse Embryo": A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor N.Y., 1986; Thomas *et al.*, 1987, *Cell* 51:503; Koller *et al.*, 1991, *Proc. Natl. Acad. Sci. USA*, 88:10730; Dorin *et al.*, 1992, *Transgenic Res.* 1:101; and Veis *et al.*, 1993, *Cell* 75:229). The ES cells that will be

20 inserted with the targeting construct are derived from an embryo or blastocyst of the same species as the developing embryo into which they are to be introduced. ES cells are typically selected for their ability to integrate into the inner cell mass and contribute to the germ line of an individual when introduced into the mammal in an embryo at the blastocyst stage of development. Thus, any ES cell line having this capability is suitable

25 for use in the practice of the present invention.

The present invention may also be used to disrupt (e.g., knockout) genes in other cell types, such as stem cells. By way of example, stem cells may be myeloid, lymphoid, or neural progenitor and precursor cells. These cells comprising a disruption or knockout of a gene may be particularly useful in the study of TSH-R gene function in individual

30 developmental pathways. Stem cells may be derived from any vertebrate species, such as mouse, rat, dog, cat, pig, rabbit, human, non-human primates and the like.

After the targeting construct has been introduced into cells, the cells in which successful gene targeting has occurred are identified. Insertion of the targeting construct into the targeted gene is typically detected by identifying cells for expression of the marker gene. In a preferred embodiment, the cells transformed with the targeting construct of the present invention are subjected to treatment with an appropriate agent that selects against cells not expressing the selectable marker. Only those cells expressing the selectable marker gene survive and/or grow under certain conditions. For example, cells that express the introduced neomycin resistance gene are resistant to the compound G418, while cells that do not express the neo gene marker are killed by G418. If the targeting construct also comprises a screening marker such as GFP, homologous recombination can be identified through screening cell colonies under a fluorescent light. Cells that have undergone homologous recombination will have deleted the GFP gene and will not fluoresce.

If a regulated positive selection method is used in identifying homologous recombination events, the targeting construct is designed so that the expression of the selectable marker gene is regulated in a manner such that expression is inhibited following random integration but is permitted (derepressed) following homologous recombination. More particularly, the transfected cells are screened for expression of the *neo* gene, which requires that (1) the cell was successfully electroporated, and (2) *lac* repressor inhibition of *neo* transcription was relieved by homologous recombination. This method allows for the identification of transfected cells and homologous recombinants to occur in one step with the addition of a single drug.

Alternatively, a positive-negative selection technique may be used to select homologous recombinants. This technique involves a process in which a first drug is added to the cell population, for example, a neomycin-like drug to select for growth of transfected cells, *i.e.* positive selection. A second drug, such as FIAU is subsequently added to kill cells that express the negative selection marker, *i.e.* negative selection. Cells that contain and express the negative selection marker are killed by a selecting agent, whereas cells that do not contain and express the negative selection marker survive. For example, cells with non-homologous insertion of the construct express HSV thymidine kinase and therefore are sensitive to the herpes drugs such as gancyclovir

17

(GANC) or FIAU (1-(2-deoxy 2-fluoro-B-D-arabinofluranosyl)-5-iodouracil). (*see, e.g.,* Mansour *et al., Nature* 336:348-352: (1988); Capecchi, *Science* 244:1288-1292, (1989); Capecchi, *Trends in Genet.* 5:70-76 (1989)).

5   Successful recombination may be identified by analyzing the DNA of the selected cells to confirm homologous recombination. Various techniques known in the art, such as PCR and/or Southern analysis may be used to confirm homologous recombination events.

Homologous recombination may also be used to disrupt genes in stem cells, and other cell types, which are not totipotent embryonic stem cells. By way of example, stem

10   cells may be myeloid, lymphoid, or neural progenitor and precursor cells. Such transgenic cells may be particularly useful in the study of TSH-R gene function in individual developmental pathways. Stem cells may be derived from any vertebrate species, such as mouse, rat, dog, cat, pig, rabbit, human, non-human primates and the like.

15   In cells that are not totipotent it may be desirable to knock out both copies of the target using methods that are known in the art. For example, cells comprising homologous recombination at a target locus that have been selected for expression of a positive selection marker (*e.g., Neo^r*) and screened for non-random integration, can be further selected for multiple copies of the selectable marker gene by exposure to elevated

20   levels of the selective agent (*e.g.,* G418). The cells are then analyzed for homozygosity at the target locus. Alternatively, a second construct can be generated with a different positive selection marker inserted between the two homologous sequences. The two constructs can be introduced into the cell either sequentially or simultaneously, followed by appropriate selection for each of the positive marker genes. The final cell is screened

25   for homologous recombination of both alleles of the target.

Production of Transgenic Animals

Selected cells are then injected into a blastocyst (or other stage of development suitable for the purposes of creating a viable animal, such as, for example, a morula) of

30   an animal (*e.g.,* a mouse) to form chimeras (*see e.g.,* Bradley, A. *in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach,* E. J. Robertson, ed., IRL, Oxford, pp.

113-152 (1987)). Alternatively, selected ES cells can be allowed to aggregate with dissociated mouse embryo cells to form the aggregation chimera. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Chimeric progeny harbouring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA. In one embodiment, chimeric progeny mice are used to generate a mouse with a heterozygous disruption in the TSH-R gene. Heterozygous transgenic mice can then be mated. It is well known in the art that typically ¼ of the offspring of such matings will have a homozygous disruption in the TSH-R gene.

The heterozygous and homozygous transgenic mice can then be compared to normal, wild-type mice to determine whether disruption of the TSH-R gene causes phenotypic changes, especially pathological changes. For example, heterozygous and homozygous mice may be evaluated for phenotypic changes by physical examination, necropsy, histology, clinical chemistry, complete blood count, body weight, organ weights, and cytological evaluation of bone marrow. Phenotypic changes may also comprise behavioral modifications or abnormalities.

In one embodiment, the phenotype (or phenotypic change) associated with a disruption in the TSH-R gene is placed into or stored in a database. Preferably, the database includes: (i) genotypic data (*e.g.*, identification of the disrupted gene) and (ii) phenotypic data (*e.g.*, phenotype(s) resulting from the gene disruption) associated with the genotypic data. The database is preferably electronic. In addition, the database is preferably combined with a search tool so that the database is searchable.

Conditional Transgenic Animals

The present invention further contemplates conditional transgenic or knockout animals, such as those produced using recombination methods. Bacteriophage P1 Cre recombinase and flp recombinase from yeast plasmids are two non-limiting examples of site-specific DNA recombinase enzymes that cleave DNA at specific target sites (lox P sites for cre recombinase and frt sites for flp recombinase) and catalyze a ligation of this DNA to a second cleaved site. A large number of suitable alternative site-specific recombinases have been described, and their genes can be used in accordance with the

19

method of the present invention. Such recombinases include the Int recombinase of bacteriophage λ (with or without Xis) (Weisberg, R. *et al.*, in *Lambda II*, (Hendrix, R., *et al.*, Eds.), Cold Spring Harbor Press, Cold Spring Harbor, NY, pp. 211-50 (1983), herein incorporated by reference); TpnI and the β-lactamase transposons (Mercier, *et al., J.*

5   *Bacteriol.*, 172:3745-57 (1990)); the Tn3 resolvase (Flanagan & Fennewald *J. Molec. Biol.*, 206:295-304 (1989); Stark, *et al., Cell*, 58:779-90 (1989)); the yeast recombinases (Matsuzaki, *et al., J. Bacteriol.*, 172:610-18 (1990)); the *B. subtilis* SpoIVC recombinase (Sato, *et al., J. Bacteriol.* 172:1092-98 (1990*)*); the Flp recombinase (Schwartz & Sadowski, *J. Molec.Biol.*, 205:647-658 (1989); Parsons, *et al., J. Biol. Chem.*,

10   265:4527-33 (1990); Golic & Lindquist, *Cell*, 59:499-509 (1989); Amin, *et al., J. Molec. Biol.*, 214:55-72 (1990*)*); the Hin recombinase (Glasgow, *et al., J. Biol. Chem.*, 264:10072-82 (1989)); immunoglobulin recombinases (Malynn, *et al., Cell*, 54:453-460 (1988)); and the Cin recombinase (Haffter & Bickle, *EMBO J.*, 7:3991-3996 (1988); Hubner, *et al., J. Molec. Biol.*, 205:493-500 (1989)), all herein incorporated by reference.

15   Such systems are discussed by Echols (*J. Biol. Chem.* 265:14697-14700 (1990)); de Villartay (*Nature*, 335:170-74 (1988)); Craig, (*Ann. Rev. Genet.*, 22:77-105 (1988)); Poyart-Salmeron, *et al.*, (*EMBO J.* 8:2425-33 (1989)); Hunger-Bertling, *et al.,(Mol Cell. Biochem.*, 92:107-16 (1990)); and Cregg & Madden (*Mol. Gen. Genet.*, 219:320-23 (1989)), all herein incorporated by reference.

20   Cre has been purified to homogeneity, and its reaction with the loxP site has been extensively characterized (Abremski & Hess *J. Mol. Biol.* 259:1509-14 (1984), herein incorporated by reference). Cre protein has a molecular weight of 35,000 and can be obtained commercially from New England Nuclear/Du Pont. The cre gene (which encodes the Cre protein) has been cloned and expressed (Abremski, *et al., Cell*

25   32:1301-11 (1983), herein incorporated by reference). The Cre protein mediates recombination between two loxP sequences (Sternberg, *et al., Cold Spring Harbor Symp. Quant. Biol.* 45:297-309 (1981)), which may be present on the same or different DNA molecule. Because the internal spacer sequence of the loxP site is asymmetrical, two loxP sites can exhibit directionality relative to one another (Hoess & Abremski *Proc.*

30   *Natl. Acad. Sci. U.S.A.* 81:1026-29 (1984*)*). Thus, when two sites on the same DNA molecule are in a directly repeated orientation, Cre will excise the DNA between the sites

(Abremski, *et al.,* *Cell* 32:1301-11 (1983)). However, if the sites are inverted with respect to each other, the DNA between them is not excised after recombination but is simply inverted. Thus, a circular DNA molecule having two loxP sites in direct orientation will recombine to produce two smaller circles, whereas circular molecules having two loxP sites in an inverted orientation simply invert the DNA sequences flanked by the loxP sites. In addition, recombinase action can result in reciprocal exchange of regions distal to the target site when targets are present on separate DNA molecules.

Recombinases have important application for characterizing gene function in knockout models. When the constructs described herein are used to disrupt TSH-R genes, a fusion transcript can be produced when insertion of the positive selection marker occurs downstream (3') of the translation initiation site of the TSH-R gene. The fusion transcript could result in some level of protein expression with unknown consequence. It has been suggested that insertion of a positive selection marker gene can affect the expression of nearby genes. These effects may make it difficult to determine gene function after a knockout event since one could not discern whether a given phenotype is associated with the inactivation of a gene, or the transcription of nearby genes. Both potential problems are solved by exploiting recombinase activity. When the positive selection marker is flanked by recombinase sites in the same orientation, the addition of the corresponding recombinase will result in the removal of the positive selection marker. In this way, effects caused by the positive selection marker or expression of fusion transcripts are avoided.

In one embodiment, purified recombinase enzyme is provided to the cell by direct microinjection. In another embodiment, recombinase is expressed from a co-transfected construct or vector in which the recombinase gene is operably linked to a functional promoter. An additional aspect of this embodiment is the use of tissue-specific or inducible recombinase constructs that allow the choice of when and where recombination occurs. One method for practicing the inducible forms of recombinase-mediated recombination involves the use of vectors that use inducible or tissue-specific promoters or other gene regulatory elements to express the desired recombinase activity. The inducible expression elements are preferably operatively positioned to allow the inducible control or activation of expression of the desired recombinase activity. Examples of such

21

inducible promoters or other gene regulatory elements include, but are not limited to, tetracycline, metallothionine, ecdysone, and other steroid-responsive promoters, rapamycin responsive promoters, and the like (No, *et al., Proc. Natl. Acad. Sci. USA,* 93:3346-51 (1996); Furth, *et al., Proc. Natl. Acad. Sci. USA,* 91:9302-6 (1994)).

5   Additional control elements that can be used include promoters requiring specific transcription factors such as viral, promoters. Vectors incorporating such promoters would only express recombinase activity in cells that express the necessary transcription factors.

10   Models for Disease

         The cell- and animal-based systems described herein can be utilized as models for diseases. Animals of any species, including, but not limited to, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, goats, and non-human primates, *e.g.*, baboons, monkeys, and chimpanzees may be used to generate disease animal models. In addition, cells from

15   humans may be used. These systems may be used in a variety of applications. Such assays may be utilized as part of screening strategies designed to identify agents, such as compounds that are capable of ameliorating disease symptoms. Thus, the animal- and cell-based models may be used to identify drugs, pharmaceuticals, therapies and interventions that may be effective in treating disease.

20   Cell-based systems may be used to identify compounds that may act to ameliorate disease symptoms. For example, such cell systems may be exposed to a compound suspected of exhibiting an ability to ameliorate disease symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of disease symptoms in the exposed cells. After exposure, the cells are examined to determine whether one or

25   more of the disease cellular phenotypes has been altered to resemble a more normal or more wild type, non-disease phenotype.

         In addition, animal-based disease systems, such as those described herein, may be used to identify compounds capable of ameliorating disease symptoms. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals,

30   therapies, and interventions that may be effective in treating a disease or other phenotypic characteristic of the animal. For example, animal models may be exposed to a compound

or agent suspected of exhibiting an ability to ameliorate disease symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of disease symptoms in the exposed animals. The response of the animals to the exposure may be monitored by assessing the reversal of disorders associated with the disease. Exposure may involve

5   treating mother animals during gestation of the model animals described herein, thereby exposing embryos or fetuses to the compound or agent that may prevent or ameliorate the disease or phenotype. Neonatal, juvenile, and adult animals can also be exposed.

More particularly, using the animal models of the invention, methods of identifying agents are provided in which such agents can be identified, on the basis of

10   their ability to affect at least one phenotype associated with a disruption in a TSH-R gene. In one embodiment, the present invention provides a method of identifying agents having an effect on TSH-R expression or function. The method includes measuring a physiological response of the animal, for example, to the agent, and comparing the physiological response of such animal to a control animal, wherein the physiological

15   response of the animal comprising a disruption in a TSH-R as compared to the control animal indicates the specificity of the agent. A "physiological response" is any biological or physical parameter of an animal that can be measured. Molecular assays (e.g., gene transcription, protein production and degradation rates), physical parameters (e.g., exercise physiology tests, measurement of various parameters of respiration,

20   measurement of heart rate or blood pressure, or measurement of bleeding time), behavioral testing, and cellular assays (e.g., immunohistochemical assays of cell surface markers, or the ability of cells to aggregate or proliferate) can be used to assess a physiological response.

The transgenic animals and cells of the present invention may be utilized as

25   models for diseases, disorders, or conditions associated with phenotypes relating to a disruption in a TSH-R gene.

The present invention provides a unique animal model for testing and developing new treatments relating to the behavioral phenotypes. Analysis of the behavioral phenotype allows for the development of an animal model useful for testing, for instance,

30   the efficacy of proposed genetic and pharmacological therapies for human genetic diseases, such as neurological, neuropsychological, or psychotic illnesses.

23

A statistical analysis of the various behaviors measured can be carried out using any conventional statistical program routinely used by those skilled in the art (such as, for example, "Analysis of Variance" or ANOVA). A "p" value of about 0.05 or less is generally considered to be statistically significant, although slightly higher p values may still be indicative of statistically significant differences. To statistically analyze abnormal behavior, a comparison is made between the behavior of a transgenic animal (or a group thereof) to the behavior of a wild-type mouse (or a group thereof), typically under certain prescribed conditions. "Abnormal behavior" as used herein refers to behavior exhibited by an animal having a disruption in the TSH-R gene, *e.g.* transgenic animal, which differs from an animal without a disruption in the TSH-R gene, *e.g.* wild-type mouse. Abnormal behavior consists of any number of standard behaviors that can be objectively measured (or observed) and compared. In the case of comparison, it is preferred that the change be statistically significant to confirm that there is indeed a meaningful behavioral difference between the knockout animal and the wild-type control animal. Examples of behaviors that may be measured or observed include, but are not limited to, ataxia, rapid limb movement, eye movement, breathing, motor activity, cognition, emotional behaviors, social behaviors, hyperactivity, hypersensitivity, anxiety, impaired learning, abnormal reward behavior, and abnormal social interaction, such as aggression.

A series of tests may be used to measure the behavioral phenotype of the animal models of the present invention, including neurological and neuropsychological tests to identify abnormal behavior. These tests may be used to measure abnormal behavior relating to, for example, learning and memory, eating, pain, aggression, sexual reproduction, anxiety, depression, schizophrenia, and drug abuse. (*see, e.g.*, Crawley & Paylor, *Hormones and Behavior* 31:197-211 (1997)).

The social interaction test involves exposing a mouse to other animals in a variety of settings. The social behaviors of the animals (*e.g.*, touching, climbing, sniffing, and mating) are subsequently evaluated. Differences in behaviors can then be statistically analyzed and compared (*see, e.g.*, S. E. File, *et al.*, *Pharmacol. Bioch. Behav.* 22:941-944 (1985); R. R. Holson, *Phys. Behav.* 37:239-247 (1986)). Examplary behavioral tests include the following.

The mouse startle response test typically involves exposing the animal to a sensory (typically auditory) stimulus and measuring the startle response of the animal (*see, e.g.,* M. A. Geyer, *et al., Brain Res. Bull.* 25:485-498 (1990); Paylor and Crawley, *Psychopharmacology* 132:169-180 (1997)). A pre-pulse inhibition test can also be used, in which the percent inhibition (from a normal startle response) is measured by "cueing" the animal first with a brief low-intensity pre-pulse prior to the startle pulse.

The electric shock test generally involves exposure to an electrified surface and measurement of subsequent behaviors such as, for example, motor activity, learning, social behaviors. The behaviors are measured and statistically analyzed using standard statistical tests. (*see, e.g.,* G. J. Kant, *et al., Pharm. Bioch. Behav.* 20:793-797 (1984); N. J. Leidenheimer, *et al., Pharmacol. Bioch. Behav.* 30:351-355 (1988)).

The tail-pinch or immobilization test involves applying pressure to the tail of the animal and/or restraining the animal's movements. Motor activity, social behavior, and cognitive behavior are examples of the areas that are measured. (*see, e.g.,* M. Bertolucci D'Angic, *et al., Neurochem.* 55:1208-1214 (1990)).

The novelty test generally comprises exposure to a novel environment and/or novel objects. The animal's motor behavior in the novel environment and/or around the novel object are measured and statistically analyzed. (*see, e.g.,* D. K. Reinstein, *et al., Pharm. Bioch. Behav.* 17:193-202 (1982); B. Poucet, *Behav. Neurosci.* 103:1009-10016 (1989); R. R. Holson, *et al., Phys. Behav.* 37:231-238 (1986)). This test may be used to detect visual processing deficiencies or defects.

The learned helplessness test involves exposure to stresses, for example, noxious stimuli, which cannot be affected by the animal's behavior. The animal's behavior can be statistically analyzed using various standard statistical tests. (*see, e.g.,* A. Leshner, *et al., Behav. Neural Biol.* 26:497-501 (1979)).

Alternatively, a tail suspension test may be used, in which the "immobile" time of the mouse is measured when suspended "upside-down" by its tail. This is a measure of whether the animal struggles, an indicator of depression. In humans, depression is believed to result from feelings of a lack of control over one's life or situation. It is believed that a depressive state can be elicited in animals by repeatedly subjecting them to aversive situations over which they have no control. A condition of "learned

25

helplessness" is eventually reached, in which the animal will stop trying to change its circumstances and simply accept its fate. Animals that stop struggling sooner are believed to be more prone to depression. Studies have shown that the administration of certain antidepressant drugs prior to testing increases the amount of time that animals

5   struggle before giving up.

The Morris water-maze test comprises learning spatial orientations in water and subsequently measuring the animal's behaviors, such as, for example, by counting the number of incorrect choices. The behaviors measured are statistically analyzed using standard statistical tests. (*see, e.g.*, E. M. Spruijt, *et al., Brain Res.* 527:192-197 (1990)).

10   Alternatively, a Y-shaped maze may be used (*see, e.g.*, McFarland, D.J., *Pharmacology, Biochemistry and Behavior* 32:723-726 (1989); Dellu, F., *et al., Neurobiology of Learning and Memory* 73:31-48 (2000)). The Y-maze is generally believed to be a test of cognitive ability. The dimensions of each arm of the Y-maze can be, for example, approximately 40 cm x 8 cm x 20 cm, although other dimensions may be

15   used. Each arm can also have, for example, sixteen equally spaced photobeams to automatically detect movement within the arms. At least two different tests can be performed using such a Y-maze. In a continuous Y-maze paradigm, mice are allowed to explore all three arms of a Y-maze for, *e.g.*, approximately 10 minutes. The animals are continuously tracked using photobeam detection grids, and the data can be used to

20   measure spontaneous alteration and positive bias behavior. Spontaneous alteration refers to the natural tendency of a "normal" animal to visit the least familiar arm of a maze. An alternation is scored when the animal makes two consecutive turns in the same direction, thus representing a sequence of visits to the least recently entered arm of the maze. Position bias determines egocentrically defined responses by measuring the animal's

25   tendency to favor turning in one direction over another. Therefore, the test can detect differences in an animal's ability to navigate on the basis of allocentric or egocentric mechanisms. The two-trial Y-maze memory test measures response to novelty and spatial memory based on a free-choice exploration paradigm. During the first trial (acquisition), the animals are allowed to freely visit two arms of the Y-maze for, *e.g.*,

30   approximately 15 minutes. The third arm is blocked off during this trial. The second trial (retrieval) is performed after an intertrial interval of, *e.g.*, approximately 2 hours.

During the retrieval trial, the blocked arm is opened and the animal is allowed access to all three arms for, *e.g.*, approximately 5 minutes. Data are collected during the retrieval trial and analyzed for the number and duration of visits to each arm. Because the three arms of the maze are virtually identical, discrimination between novelty and familiarity is

5   dependent on "environmental" spatial cues around the room relative to the position of each arm. Changes in arm entry and duration of time spent in the novel arm in a transgenic animal model may be indicative of a role of that gene in mediating novelty and recognition processes.

The passive avoidance or shuttle box test generally involves exposure to two or

10   more environments, one of which is noxious, providing a choice to be learned by the animal. Behavioral measures include, for example, response latency, number of correct responses, and consistency of response. (*see, e.g.*, R. Ader, *et al.*, *Psychon. Sci.* 26:125-128 (1972); R. R. Holson, *Phys. Behav.* 37:221-230 (1986)). Alternatively, a zero-maze can be used. In a zero-maze, the animals can, for example, be placed in a closed quadrant

15   of an elevated annular platform having, *e.g.*, 2 open and 2 closed quadrants, and are allowed to explore for approximately 5 minutes. This paradigm exploits an approach-avoidance conflict between normal exploratory activity and an aversion to open spaces in rodents. This test measures anxiety levels and can be used to evaluate the effectiveness of anti-anxiolytic drugs. The time spent in open quadrants versus closed quadrants may

20   be recorded automatically, with, for example, the placement of photobeams at each transition site.

The food avoidance test involves exposure to novel food and objectively measuring, for example, food intake and intake latency. The behaviors measured are statistically analyzed using standard statistical tests. (*see, e.g.*, B. A. Campbell, *et al.*, *J.*

25   *Comp. Physiol. Psychol.* 67:15-22 (1969)).

The elevated plus-maze test comprises exposure to a maze, without sides, on a platform, the animal's behavior is objectively measured by counting the number of maze entries and maze learning. The behavior is statistically analyzed using standard statistical tests. (*see, e.g.*, H. A. Baldwin, *et al.*, *Brain Res. Bull*, 20:603-606 (1988)).

30   The stimulant-induced hyperactivity test involves injection of stimulant drugs (*e.g.*, amphetamines, cocaine, PCP, and the like), and objectively measuring, for example,

27

motor activity, social interactions, cognitive behavior. The animal's behaviors are statistically analyzed using standard statistical tests. (*see, e.g.*, P. B. S. Clarke, *et al.*, *Psychopharmacology* 96:511-520 (1988); P. Kuczenski, *et al.*, *J. Neuroscience* 11:2703-2712 (1991)).

5   The self-stimulation test generally comprises providing the mouse with the opportunity to regulate electrical and/or chemical stimuli to its own brain. Behavior is measured by frequency and pattern of self-stimulation. Such behaviors are statistically analyzed using standard statistical tests. (*see, e.g.*, S. Nassif, *et al.*, *Brain Res.*, 332:247-257 (1985); W. L. Isaac, *et al.*, *Behav. Neurosci.* 103:345-355 (1989)).

10   The reward test involves shaping a variety of behaviors, *e.g.*, motor, cognitive, and social, measuring, for example, rapidity and reliability of behavioral change, and statistically analyzing the behaviors measured. (*see, e.g.*, L. E. Jarrard, *et al.*, Exp. Brain Res. 61:519-530 (1986)).

The DRL (differential reinforcement to low rates of responding) performance test

15   involves exposure to intermittent reward paradigms and measuring the number of proper responses, *e.g.*, lever pressing. Such behavior is statistically analyzed using standard statistical tests. (*see, e.g.*, J. D. Sinden, *et al.*, Behav. Neurosci. 100:320-329 (1986); V. Nalwa, *et al.*, Behav Brain Res. 17:73-76 (1985); and A. J. Nonneman, *et al.*, J. Comp. Physiol. Psych. 95:588-602 (1981)).

20   The spatial learning test involves exposure to a complex novel environment, measuring the rapidity and extent of spatial learning, and statistically analyzing the behaviors measured. (*see, e.g.*, N. Pitsikas, *et al.*, Pharm. Bioch. Behav. 38:931-934 (1991); B. poucet, *et al.*, Brain Res. 37:269-280 (1990); D. Christie, *et al.*, Brain Res. 37:263-268 (1990); and F. Van Haaren, *et al.*, Behav. Neurosci. 102:481-488 (1988)).

25   Alternatively, an open-field (of) test may be used, in which the greater distance traveled for a given amount of time is a measure of the activity level and anxiety of the animal. When the open field is a novel environment, it is believed that an approach-avoidance situation is created, in which the animal is "torn" between the drive to explore and the drive to protect itself. Because the chamber is lighted and has no places to hide other

30   than the corners, it is expected that a "normal" mouse will spend more time in the corners and around the periphery than it will in the center where there is no place to hide.

"Normal" mice will, however, venture into the central regions as they explore more and more of the chamber. It can then be extrapolated that especially anxious mice will spend most of their time in the corners, with relatively little or no exploration of the central region, whereas bold (*i.e.*, less anxious) mice will travel a greater distance, showing little

5   preference for the periphery versus the central region.

The visual, somatosensory and auditory neglect tests generally comprise exposure to a sensory stimulus, objectively measuring, for example, orientating responses, and statistically analyzing the behaviors measured. (*see*, *e.g.*, J. M. Vargo, *et al.*, Exp. Neurol. 102:199-209 (1988)).

10   The consummatory behavior test generally comprises feeding and drinking, and objectively measuring quantity of consumption. The behavior measured is statistically analyzed using standard statistical tests. (*see*, *e.g.*, P. J. Fletcher, *et al.*, Psychopharmacol. 102:301-308 (1990); M. G. Corda, *et al.*,, Proc. Nat'l Acad. Sci. USA 80:2072-2076 (1983)).

15   A visual discrimination test can also be used to evaluate the visual processing of an animal. One or two similar objects are placed in an open field and the animal is allowed to explore for about 5-10 minutes.. The time spent exploring each object (proximity to, *i.e.*, movement within, *e.g.*, about 3-5 cm of the object is considered exploration of an object) is recorded. The animal is then removed from the open field,

20   and the objects are replaced by a similar object and a novel object. The animal is returned to the open field and the percent time spent exploring the novel object over the old object is measured (again, over about a 5-10 minute span). "Normal" animals will typically spend a higher percentage of time exploring the novel object rather than the old object. If a delay is imposed between sampling and testing, the memory task becomes

25   more hippocampal-dependent. If no delay is imposed, the task is more based on simple visual discrimination. This test can also be used for olfactory discrimination, in which the objects (preferably, simple blocks) can be sprayed or otherwise treated to hold an odor. This test can also be used to determine if the animal can make gustatory discriminations; animals that return to the previously eaten food instead of novel food

30   exhibit gustatory neophobia.

A hot plate analgesia test can be used to evaluate an animal's sensitivity to heat or painful stimuli. For example, a mouse can be placed on an approximately 55°C hot plate and the mouse's response latency (*e.g.*, time to pick up and lick a hind paw) can be recorded. These responses are not reflexes, but rather "higher" responses requiring cortical involvement. This test may be used to evaluate a nociceptive disorder.

An accelerating rotarod test may be used to measure coordination and balance in mice. Animals can be, for example, placed on a rod that acts like a rotating treadmill (or rolling log). The rotarod can be made to rotate slowly at first and then progressively faster until it reaches a speed of, *e.g.*, approximately 60 rpm. The mice must continually reposition themselves in order to avoid falling off. The animals are preferably tested in at least three trials, a minimum of 20 minutes apart. Those mice that are able to stay on the rod the longest are believed to have better coordination and balance.

A metrazol administration test can be used to screen animals for varying susceptibilies to seizures or similar events. For example, a 5mg/ml solution of metrazol can be infused through the tail vein of a mouse at a rate of, *e.g.*, approximately 0.375 ml/min. The infusion will cause all mice to experience seizures, followed by death. Those mice that enter the seizure stage the soonest are believed to be more prone to seizures. Four distinct physiological stages can be recorded: soon after the start of infusion, the mice will exhibit a noticeable "twitch", followed by a series of seizures, ending in a final tensing of the body known as "tonic extension", which is followed by death.

TSH-R Gene Products

The present invention further contemplates use of the TSH-R gene sequence to produce TSH-R gene products. TSH-R gene products may include proteins that represent functionally equivalent gene products. Such an equivalent gene product may contain deletions, additions or substitutions of amino acid residues within the amino acid sequence encoded by the gene sequences described herein, but which result in a silent change, thus producing a functionally equivalent TSH-R gene product. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved.

For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine;

5   and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Functionally equivalent", as utilized herein, refers to a protein capable of exhibiting a substantially similar *in vivo* activity as the endogenous gene products encoded by the TSH-R gene sequences. Alternatively, when utilized as part of an assay, "functionally equivalent" may refer to peptides capable of interacting with other cellular or

10   extracellular molecules in a manner substantially similar to the way in which the corresponding portion of the endogenous gene product would.

Other protein products useful according to the methods of the invention are peptides derived from or based on the TSH-R gene produced by recombinant or synthetic means (derived peptides).

15   TSH-R gene products may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the gene polypeptides and peptides of the invention by expressing nucleic acid encoding gene sequences are described herein. Methods that are well known to those skilled in the art can be used to construct expression vectors containing gene protein coding sequences and appropriate

20   transcriptional/translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques and *in vivo* recombination/genetic recombination (*see, e.g.*, Sambrook, *et al.*, 1989, *supra*, and Ausubel, *et al.*, 1989, supra). Alternatively, RNA capable of encoding gene protein sequences may be chemically synthesized using, for example, automated synthesizers (*see, e.g.* Oligonucleotide

25   Synthesis: A Practical Approach, Gait, M. J. ed., IRL Press, Oxford (1984)).

A variety of host-expression vector systems may be utilized to express the gene coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells that may, when transformed or transfected with the appropriate

30   nucleotide coding sequences, exhibit the gene protein of the invention *in situ*. These include but are not limited to microorganisms such as bacteria (*e.g., E. coli, B. subtilis*)

31

transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing gene protein coding sequences; yeast (*e.g. Saccharomyces, Pichia*) transformed with recombinant yeast expression vectors containing the gene protein coding sequences; insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus) containing the gene protein coding sequences; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing gene protein coding sequences; or mammalian cell systems (*e.g.* COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (*e.g.*, metallothionine promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5 K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the gene protein being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of antibodies or to screen peptide libraries, for example, vectors that direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther *et al., EMBO J.*, 2:1791-94 (1983)), in which the gene protein coding sequence may be ligated individually into the vector in frame with the *lac Z* coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, *Nucleic Acids Res.*, 13:3101-09 (1985); Van Heeke *et al., J. Biol. Chem.*, 264:5503-9 (1989)); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned TSH-R gene protein can be released from the GST moiety.

In a preferred embodiment, full length cDNA sequences are appended with in-frame Bam HI sites at the amino terminus and Eco RI sites at the carboxyl terminus using standard PCR methodologies (Innis, *et al.* (eds) PCR Protocols: A Guide to Methods and

Applications, Academic Press, San Diego (1990)) and ligated into the pGEX-2TK vector (Pharmacia, Uppsala, Sweden). The resulting cDNA construct contains a kinase recognition site at the amino terminus for radioactive labeling and glutathione S-transferase sequences at the carboxyl terminus for affinity purification (Nilsson, *et al.*,

5   *EMBO J.*, 4: 1075-80 (1985); Zabeau *et al.*, *EMBO J.*, 1: 1217-24 (1982)).

In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in Spodoptera frugiperda cells. The gene coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV

10   promoter (for example the polyhedrin promoter). Successful insertion of gene coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed (*see, e.g.*, Smith, *et al.*, *J. Virol.*

15   46: 584-93 (1983); U.S. Patent No. 4,745,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the gene coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene may

20   then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing gene protein in infected hosts. (*e.g., see* Logan *et al.*, *Proc. Natl. Acad. Sci. USA*, 81:3655-59 (1984)). Specific initiation signals may also be required for efficient translation of inserted gene coding sequences.

25   These signals include the ATG initiation codon and adjacent sequences. In cases where an entire gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the gene coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon,

30   must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These

exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (*see* Bitter, *et al., Methods in Enzymol.*, 153:516-44 (1987)).

5   In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (*e.g.*, glycosylation) and processing (*e.g.*, cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and

10   modification of proteins. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa,

15   COS, MDCK, 293, 3T3, WI38, etc.

        For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express the gene protein may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements

20   (*e.g.*, promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells that stably integrate the plasmid into their

25   chromosomes and grow, to form foci, which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines that express the gene protein. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the gene protein.

        In a preferred embodiment, timing and/or quantity of expression of the

30   recombinant protein can be controlled using an inducible expression construct. Inducible constructs and systems for inducible expression of recombinant proteins will be well

known to those skilled in the art. Examples of such inducible promoters or other gene regulatory elements include, but are not limited to, tetracycline, metallothionine, ecdysone, and other steroid-responsive promoters, rapamycin responsive promoters, and the like (No, *et al.*, *Proc. Natl. Acad. Sci. USA*, 93:3346-51 (1996); Furth, *et al.*, *Proc.*

5   *Natl. Acad. Sci. USA*, 91:9302-6 (1994)). Additional control elements that can be used include promoters requiring specific transcription factors such as viral, particularly HIV, promoters. In one in embodiment, a Tet inducible gene expression system is utilized. (Gossen *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:5547-51 (1992); Gossen, *et al.*, *Science*, 268:1766-69 (1995)). Tet Expression Systems are based on two regulatory elements

10   derived from the tetracycline-resistance operon of the *E. coli* Tn10 transposon—the tetracycline repressor protein (TetR) and the tetracycline operator sequence (*tetO*) to which TetR binds. Using such a system, expression of the recombinant protein is placed under the control of the *tetO* operator sequence and transfected or transformed into a host cell. In the presence of TetR, which is co-transfected into the host cell, expression of the

15   recombinant protein is repressed due to binding of the TetR protein to the *tetO* regulatory element. High-level, regulated gene expression can then be induced in response to varying concentrations of tetracycline (Tc) or Tc derivatives such as doxycycline (Dox), which compete with *tetO* elements for binding to TetR. Constructs and materials for tet inducible gene expression are available commercially from CLONTECH Laboratories,

20   Inc., Palo Alto, CA.

When used as a component in an assay system, the gene protein may be labeled, either directly or indirectly, to facilitate detection of a complex formed between the gene protein and a test substance. Any of a variety of suitable labeling systems may be used including but not limited to radioisotopes such as ^{125}I; enzyme labeling systems that

25   generate a detectable calorimetric signal or light when exposed to substrate; and fluorescent labels. Where recombinant DNA technology is used to produce the gene protein for such assay systems, it may be advantageous to engineer fusion proteins that can facilitate labeling, immobilization and/or detection.

Indirect labeling involves the use of a protein, such as a labeled antibody, which

30   specifically binds to the gene product. Such antibodies include but are not limited to

polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by a Fab expression library.

<u>Production of Antibodies</u>

5   Described herein are methods for the production of antibodies capable of specifically recognizing one or more epitopes. Such antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding

10   fragments of any of the above. Such antibodies may be used, for example, in the detection of a TSH-R gene in a biological sample, or, alternatively, as a method for the inhibition of abnormal TSH-R gene activity. Thus, such antibodies may be utilized as part of disease treatment methods, and/or may be used as part of diagnostic techniques whereby patients may be tested for abnormal levels of TSH-R gene proteins, or for the

15   presence of abnormal forms of such proteins.

For the production of antibodies, various host animals may be immunized by injection with the TSH-R gene, its expression product or a portion thereof. Such host animals may include but are not limited to rabbits, mice, rats, goats and chickens, to name but a few. Various adjuvants may be used to increase the immunological response,

20   depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (*bacille Calmette-Guerin*) and *Corynebacterium parvum*.

25   Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as TSH-R gene product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described above, may be immunized by injection with gene product supplemented with adjuvants as also described above.

30   Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by any technique that provides for the production of

antibody molecules by continuous cell lines in culture. These include, but are not limited to the hybridoma technique of Köhler and Milstein, *Nature*, 256:495-7 (1975); and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor, *et al.*, *Immunology Today*, 4:72 (1983); Cote, *et al.*, *Proc. Natl. Acad. Sci. USA*, 80:2026-30

5   (1983)), and the EBV-hybridoma technique (Cole, *et al.*, in Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., New York, pp. 77-96 (1985)). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated *in vitro* or *in vivo*. Production of high titers of mAbs *in vivo* makes this the presently preferred

10   method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison, *et al.*, *Proc. Natl. Acad. Sci.*, 81:6851-6855 (1984); Takeda, *et al.*, *Nature*, 314:452-54 (1985)) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate

15   biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778; Bird, *Science* 242:423-26 (1988); Huston, *et al.*, *Proc. Natl.*

20   *Acad. Sci. USA*, 85:5879-83 (1988); and Ward, *et al.*, *Nature*, 334:544-46 (1989)) can be adapted to produce gene-single chain antibodies. Single chain antibodies are typically formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments that recognize specific epitopes may be generated by known

25   techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragments that can be produced by pepsin digestion of the antibody molecule and the Fab fragments that can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed (Huse, *et al.*, *Science*, 246:1275-81 (1989)) to allow rapid and easy identification of monoclonal Fab

30   fragments with the desired specificity.

Screening Methods

The present invention may be employed in a process for screening for agents such as agonists, *i.e.* agents that bind to and activate TSH-R polypeptides, or antagonists, *i.e.* inhibit the activity or interaction of TSH-R polypeptides with its ligand. Thus,

5   polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures as known in the art. Any methods routinely used to identify and screen for agents that can modulate receptors may be used in accordance with the present invention.

10   The present invention provides methods for identifying and screening for agents that modulate TSH-R expression or function. More particularly, cells that contain and express TSH-R gene sequences may be used to screen for therapeutic agents. Such cells may include non-recombinant monocyte cell lines, such as U937 (ATCC# CRL-1593), THP-1 (ATCC# TIB-202), and P388D1 (ATCC# TIB-63); endothelial cells such as

15   HUVEC's and bovine aortic endothelial cells (BAEC's); as well as generic mammalian cell lines such as HeLa cells and COS cells, *e.g.*, COS-7 (ATCC# CRL-1651). Further, such cells may include recombinant, transgenic cell lines. For example, the transgenic mice of the invention may be used to generate cell lines, containing one or more cell types involved in a disease, that can be used as cell culture models for that disorder.

20   While cells, tissues, and primary cultures derived from the disease transgenic animals of the invention may be utilized, the generation of continuous cell lines is preferred. For examples of techniques that may be used to derive a continuous cell line from the transgenic animals, see Small, *et al.*, *Mol. Cell Biol.*, 5:642-48 (1985).

TSH-R gene sequences may be introduced into, and overexpressed in, the genome

25   of the cell of interest. In order to overexpress a TSH-R gene sequence, the coding portion of the TSH-R gene sequence may be ligated to a regulatory sequence that is capable of driving gene expression in the cell type of interest. Such regulatory regions will be well known to those of skill in the art, and may be utilized in the absence of undue experimentation. TSH-R gene sequences may also be disrupted or underexpressed. Cells

30   having TSH-R gene disruptions or underexpressed TSH-R gene sequences may be used,

for example, to screen for agents capable of affecting alternative pathways that compensate for any loss of function attributable to the disruption or underexpression.

In vitro systems may be designed to identify compounds capable of binding the TSH-R gene products. Such compounds may include, but are not limited to, peptides

5   made of D-and/or L-configuration amino acids (in, for example, the form of random peptide libraries; (*see e.g.*, Lam, *et al.*, *Nature*, 354:82-4 (1991)), phosphopeptides (in, for example, the form of random or partially degenerate, directed phosphopeptide libraries; *see, e.g.*, Songyang, *et al.*, *Cell*, 72:767-78 (1993)), antibodies, and small organic or inorganic molecules. Compounds identified may be useful, for example, in

10   modulating the activity of TSH-R gene proteins, preferably mutant TSH-R gene proteins; elaborating the biological function of the TSH-R gene protein; or screening for compounds that disrupt normal TSH-R gene interactions or themselves disrupt such interactions.

The principle of the assays used to identify compounds that bind to the TSH-R

15   gene protein involves preparing a reaction mixture of the TSH-R gene protein and the test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex that can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways. For example, one method to conduct such an assay would involve anchoring the TSH-R gene protein or the

20   test substance onto a solid phase and detecting target protein/test substance complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, the TSH-R gene protein may be anchored onto a solid surface, and the test compound, which is not anchored, may be labeled, either directly or indirectly.

In practice, microtitre plates are conveniently utilized. The anchored component

25   may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein may be used to anchor the protein to the solid surface. The surfaces may be prepared in advance and stored.

30   In order to conduct the assay, the nonimmobilized component is added to the coated surface containing the anchored component. After the reaction is complete,

39

unreacted components are removed (*e.g.*, by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously nonimmobilized component is pre-labeled, the detection of label

5   immobilized on the surface indicates that complexes were formed. Where the previously nonimmobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the previously nonimmobilized component (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

10   Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized antibody specific for TSH-R gene product or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

15   Compounds that are shown to bind to a particular TSH-R gene product through one of the methods described above can be further tested for their ability to elicit a biochemical response from the TSH-R gene protein. Agonists, antagonists and/or inhibitors of the expression product can be identified utilizing assays well known in the art.

20

Antisense, Ribozymes, and Antibodies

Other agents that may be used as therapeutics include the TSH-R gene, its expression product(s) and functional fragments thereof. Additionally, agents that reduce or inhibit mutant TSH-R gene activity may be used to ameliorate disease symptoms.

25   Such agents include antisense, ribozyme, and triple helix molecules. Techniques for the production and use of such molecules are well known to those of skill in the art.

Anti-sense RNA and DNA molecules act to directly block the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. With respect to antisense DNA, oligodeoxyribonucleotides derived from the translation

30   initiation site, *e.g.*, between the -10 and +10 regions of the TSH-R gene nucleotide sequence of interest, are preferred.

40

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage. The composition of ribozyme molecules must include one or

5   more sequences complementary to the TSH-R gene mRNA, and must include the well known catalytic sequence responsible for mRNA cleavage. For this sequence, see U.S. Patent No. 5,093,246, which is incorporated by reference herein in its entirety. As such within the scope of the invention are engineered hammerhead motif ribozyme molecules that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences

10   encoding TSH-R gene proteins.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the molecule of interest for ribozyme cleavage sites that include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the TSH-R gene

15   containing the cleavage site may be evaluated for predicted structural features, such as secondary structure, that may render the oligonucleotide sequence unsuitable. The suitability of candidate sequences may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

20   Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription should be single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides must be designed to promote triple helix formation via Hoogsteen base pairing rules, which generally require sizeable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may

25   be pyrimidine-based, which will result in TAT and CGC triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, containing a stretch of G residues. These molecules will form a

30   triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the

41

purine residues are located on a single strand of the targeted duplex, resulting in GGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule.

5   Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

It is possible that the antisense, ribozyme, and/or triple helix molecules described herein may reduce or inhibit the transcription (triple helix) and/or translation (antisense,

10   ribozyme) of mRNA produced by both normal and mutant TSH-R gene alleles. In order to ensure that substantially normal levels of TSH-R gene activity are maintained, nucleic acid molecules that encode and express TSH-R polypeptides exhibiting normal activity may be introduced into cells that do not contain sequences susceptible to whatever antisense, ribozyme, or triple helix treatments are being utilized. Alternatively, it may be

15   preferable to coadminister normal TSH-R protein into the cell or tissue in order to maintain the requisite level of cellular or tissue TSH-R gene activity.

Anti-sense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing

20   oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors that incorporate suitable RNA polymerase promoters such as the T7 or SP6

25   polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Various well-known modifications to the DNA molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include

30   but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5' and/or 3' ends of the molecule or the use of

42

phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

Antibodies that are both specific for TSH-R protein, and in particular, the mutant TSH-R protein, and interfere with its activity may be used to inhibit mutant TSH-R

5   function. Such antibodies may be generated against the proteins themselves or against peptides corresponding to portions of the proteins using standard techniques known in the art and as also described herein. Such antibodies include but are not limited to polyclonal, monoclonal, Fab fragments, single chain antibodies, chimeric antibodies, etc.

In instances where the TSH-R protein is intracellular and whole antibodies are

10   used, internalizing antibodies may be preferred. However, lipofectin liposomes may be used to deliver the antibody or a fragment of the Fab region that binds to the TSH-R gene epitope into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target or expanded target protein's binding domain is preferred. For example, peptides having an amino acid sequence corresponding to the domain of the

15   variable region of the antibody that binds to the TSH-R protein may be used. Such peptides may be synthesized chemically or produced via recombinant DNA technology using methods well known in the art (see, e.g., Creighton, Proteins: Structures and Molecular Principles (1984) W.H. Freeman, New York 1983, supra; and Sambrook, et al., 1989, supra). Alternatively, single chain neutralizing antibodies that bind to

20   intracellular TSH-R gene epitopes may also be administered. Such single chain antibodies may be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population by utilizing, for example, techniques such as those described in Marasco, et al., Proc. Natl. Acad. Sci. USA, 90:7889-93 (1993).

25   RNA sequences encoding TSH-R protein may be directly administered to a patient exhibiting disease symptoms, at a concentration sufficient to produce a level of TSH-R gene protein such that disease symptoms are ameliorated. Patients may be treated by gene replacement therapy. One or more copies of a normal TSH-R gene, or a portion of the gene that directs the production of a normal TSH-R protein with TSH-R gene

30   function, may be inserted into cells using vectors that include, but are not limited to adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles

that introduce DNA into cells, such as liposomes. Additionally, techniques such as those described above may be utilized for the introduction of normal TSH-R gene sequences into human cells.

Cells, preferably, autologous cells, containing normal TSH-R gene expressing gene sequences may then be introduced or reintroduced into the patient at positions that allow for the amelioration of disease symptoms.

Pharmaceutical Compositions, Effective Dosages, and Routes of Administration

The identified compounds that inhibit target mutant gene expression, synthesis and/or activity can be administered to a patient at therapeutically effective doses to treat or ameliorate the disease. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of the disease.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds that exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (i.e., the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in

humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral, topical, subcutaneous, intraperitoneal, intraveneous, intrapleural, intraoccular, intraarterial, or rectal administration. It is also contemplated that pharmaceutical compositions may be administered with other products that potentiate the activity of the compound and optionally, may include other therapeutic ingredients.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (*e.g.*, pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (*e.g.*, lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (*e.g.*, magnesium stearate, talc or silica); disintegrants (*e.g.*, potato starch or sodium starch glycolate); or wetting agents (*e.g.*, sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (*e.g.*, sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (*e.g.*, lecithin or acacia); non-aqueous vehicles (*e.g.*, almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (*e.g.*, methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

45

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon

5   dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of *e.g.* gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

     The compounds may be formulated for parenteral administration by injection,

10   *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may

15   be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

     The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. Oral ingestion is possibly the easiest method of taking

20   any medication. Such a route of administration, is generally simple and straightforward and is frequently the least inconvenient or unpleasant route of administration from the patient's point of view. However, this involves passing the material through the stomach, which is a hostile environment for many materials, including proteins and other biologically active compositions. As the acidic, hydrolytic and proteolytic environment

25   of the stomach has evolved efficiently to digest proteinaceous materials into amino acids and oligopeptides for subsequent anabolism, it is hardly surprising that very little or any of a wide variety of biologically active proteinaceous material, if simply taken orally, would survive its passage through the stomach to be taken up by the body in the small intestine. The result, is that many proteinaceous medicaments must be taken in through

30   another method, such as parenterally, often by subcutaneous, intramuscular or intravenous injection.

Pharmaceutical compositions may also include various buffers (*e.g.*, Tris, acetate, phosphate), solubilizers (*e.g.*, Tween, Polysorbate), carriers such as human serum albumin, preservatives (thimerosol, benzyl alcohol) and anti-oxidants such as ascorbic acid in order to stabilize pharmaceutical activity. The stabilizing agent may be a

5 detergent, such as tween-20, tween-80, NP-40 or Triton X-100. EBP may also be incorporated into particulate preparations of polymeric compounds for controlled delivery to a patient over an extended period of time. A more extensive survey of components in pharmaceutical compositions is found in Remington's Pharmaceutical Sciences, 18th ed., A. R. Gennaro, ed., Mack Publishing, Easton, Pa. (1990).

10 In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion

15 exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

The compositions may, if desired, be presented in a pack or dispenser device that may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or

20 dispenser device may be accompanied by instructions for administration.

Diagnostics

A variety of methods may be employed to diagnose disease conditions associated with the TSH-R gene. Specifically, reagents may be used, for example, for the detection

25 of the presence of TSH-R gene mutations, or the detection of either over or under expression of TSH-R gene mRNA.

According to the diagnostic and prognostic method of the present invention, alteration of the wild-type TSH-R gene locus is detected. In addition, the method can be performed by detecting the wild-type TSH-R gene locus and confirming the lack of a

30 predisposition or neoplasia. "Alteration of a wild-type gene" encompasses all forms of mutations including deletions, insertions and point mutations in the coding and

noncoding regions. Deletions may be of the entire gene or only a portion of the gene. Point mutations may result in stop codons, frameshift mutations or amino acid substitutions. Somatic mutations are those that occur only in certain tissues, *e.g.*, in tumor tissue, and are not inherited in the germline. Germline mutations can be found in any of a body's tissues and are inherited. If only a single allele is somatically mutated, an early neoplastic state may be indicated. However, if both alleles are mutated, then a late neoplastic state may be indicated. The finding of gene mutations thus provides both diagnostic and prognostic information. a TSH-R gene allele that is not deleted (*e.g.*, that found on the sister chromosome to a chromosome carrying a TSH-R gene deletion) can be screened for other mutations, such as insertions, small deletions, and point mutations. Mutations found in tumor tissues may be linked to decreased expression of the TSH-R gene product. However, mutations leading to non-functional gene products may also be linked to a cancerous state. Point mutational events may occur in regulatory regions, such as in the promoter of the gene, leading to loss or diminution of expression of the mRNA. Point mutations may also abolish proper RNA processing, leading to loss of expression of the TSH-R gene product, or a decrease in mRNA stability or translation efficiency.

One test available for detecting mutations in a candidate locus is to directly compare genomic target sequences from cancer patients with those from a control population. Alternatively, one could sequence messenger RNA after amplification, *e.g.*, by PCR, thereby eliminating the necessity of determining the exon structure of the candidate gene. Mutations from cancer patients falling outside the coding region of the TSH-R gene can be detected by examining the non-coding regions, such as introns and regulatory sequences near or within the TSH-R gene. An early indication that mutations in noncoding regions are important may come from Northern blot experiments that reveal messenger RNA molecules of abnormal size or abundance in cancer patients as compared to control individuals.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific gene nucleic acid or anti-gene antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical

48

settings, to diagnose patients exhibiting disease symptoms or at risk for developing disease.

Any cell type or tissue, including brain, cortex, subcortical region, cerebellum, brainstem, olfactory bulb, spinal cord, eye, Harderian gland, heart, lung, liver, pancreas, kidney, spleen, thymus, lymph nodes, bone marrow, skin, gallbladder, urinary bladder, pituitary gland, adrenal gland, salivary gland, skeletal muscle, tongue, stomach, small intestine, large intestine, cecum, testis, epididymis, seminal vesicle, coagulating gland, prostate gland, ovary, uterus and white fat, in which the gene is expressed may be utilized in the diagnostics described below.

DNA or RNA from the cell type or tissue to be analyzed may easily be isolated using procedures that are well known to those in the art. Diagnostic procedures may also be performed *in situ* directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents may be used as probes and/or primers for such *in situ* procedures (*see*, for example, Nuovo, PCR *In Situ* Hybridization: Protocols and Applications, Raven Press, N.Y. (1992)).

Gene nucleotide sequences, either RNA or DNA, may, for example, be used in hybridization or amplification assays of biological samples to detect disease-related gene structures and expression. Such assays may include, but are not limited to, Southern or Northern analyses, restriction fragment length polymorphism assays, single stranded conformational polymorphism analyses, *in situ* hybridization assays, and polymerase chain reaction analyses. Such analyses may reveal both quantitative aspects of the expression pattern of the gene, and qualitative aspects of the gene expression and/or gene composition. That is, such aspects may include, for example, point mutations, insertions, deletions, chromosomal rearrangements, and/or activation or inactivation of gene expression.

Preferred diagnostic methods for the detection of gene-specific nucleic acid molecules may involve for example, contacting and incubating nucleic acids, derived from the cell type or tissue being analyzed, with one or more labeled nucleic acid reagents under conditions favorable for the specific annealing of these reagents to their complementary sequences within the nucleic acid molecule of interest. Preferably, the

lengths of these nucleic acid reagents are at least 9 to 30 nucleotides. After incubation, all non-annealed nucleic acids are removed from the nucleic acid:fingerprint molecule hybrid. The presence of nucleic acids from the fingerprint tissue that have hybridized, if any such molecules exist, is then detected. Using such a detection scheme, the nucleic acid from the tissue or cell type of interest may be immobilized, for example, to a solid support such as a membrane, or a plastic surface such as that on a microtitre plate or polystyrene beads. In this case, after incubation, non-annealed, labeled nucleic acid reagents are easily removed. Detection of the remaining, annealed, labeled nucleic acid reagents is accomplished using standard techniques well-known to those in the art.

Alternative diagnostic methods for the detection of gene-specific nucleic acid molecules may involve their amplification, *e.g.*, by PCR (the experimental embodiment set forth in Mullis U.S. Patent No. 4,683,202 (1987)), ligase chain reaction (Barany, *Proc. Natl. Acad. Sci. USA*, 88:189-93 (1991)), self sustained sequence replication (Guatelli, *et al.*, *Proc. Natl. Acad. Sci.* USA, 87:1874-78 (1990)), transcriptional amplification system (Kwoh, *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:1173-77 (1989)), Q-Beta Replicase (Lizardi *et al.*, *Bio/Technology*, 6:1197 (1988)), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In one embodiment of such a detection scheme, a cDNA molecule is obtained from an RNA molecule of interest (*e.g.*, by reverse transcription of the RNA molecule into cDNA). Cell types or tissues from which such RNA may be isolated include any tissue in which wild-type fingerprint gene is known to be expressed, including, but not limited, to brain, cortex, subcortical region, cerebellum, brainstem, olfactory bulb, spinal cord, eye, Harderian gland, heart, lung, liver, pancreas, kidney, spleen, thymus, lymph nodes, bone marrow, skin, gallbladder, urinary bladder, pituitary gland, adrenal gland, salivary gland, skeletal muscle, tongue, stomach, small intestine, large intestine, cecum, testis, epididymis, seminal vesicle, coagulating gland, prostate gland, ovary, uterus and white fat. A sequence within the cDNA is then used as the template for a nucleic acid amplification reaction, such as a PCR amplification reaction, or the like. The nucleic

acid reagents used as synthesis initiation reagents (*e.g.*, primers) in the reverse transcription and nucleic acid amplification steps of this method may be chosen from among the gene nucleic acid reagents described herein. The preferred lengths of such nucleic acid reagents are at least 15-30 nucleotides. For detection of the amplified

5   product, the nucleic acid amplification may be performed using radioactively or non-radioactively labeled nucleotides. Alternatively, enough amplified product may be made such that the product may be visualized by standard ethidium bromide staining or by utilizing any other suitable nucleic acid staining method.

Antibodies directed against wild type or mutant gene peptides may also be used as

10   disease diagnostics and prognostics. Such diagnostic methods, may be used to detect abnormalities in the level of gene protein expression, or abnormalities in the structure and/or tissue, cellular, or subcellular location of fingerprint gene protein. Structural differences may include, for example, differences in the size, electronegativity, or antigenicity of the mutant fingerprint gene protein relative to the normal fingerprint gene

15   protein.

Protein from the tissue or cell type to be analyzed may easily be detected or isolated using techniques that are well known to those of skill in the art, including but not limited to western blot analysis. For a detailed explanation of methods for carrying out western blot analysis, see Sambrook, *et al.* (1989) *supra*, at Chapter 18. The protein

20   detection and isolation methods employed herein may also be such as those described in Harlow and Lane, for example, (Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988)).

Preferred diagnostic methods for the detection of wild-type or mutant gene peptide molecules may involve, for example, immunoassays wherein fingerprint gene

25   peptides are detected by their interaction with an anti-fingerprint gene-specific peptide antibody.

For example, antibodies, or fragments of antibodies useful in the present invention may be used to quantitatively or qualitatively detect the presence of wild type or mutant gene peptides. This can be accomplished, for example, by

30   immunofluorescence techniques employing a fluorescently labeled antibody (see below) coupled with light microscopic, flow cytometric, or fluorimetric detection. Such

51

techniques are especially preferred if the fingerprint gene peptides are expressed on the cell surface.

The antibodies (or fragments thereof) useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for *in situ* detection of fingerprint gene peptides. *In situ* detection may be accomplished by removing a histological specimen from a patient, and applying thereto a labeled antibody of the present invention. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of the fingerprint gene peptides, but also their distribution in the examined tissue. Using the present invention, those of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such *in situ* detection.

Immunoassays for wild type, mutant, or expanded fingerprint gene peptides typically comprise incubating a biological sample, such as a biological fluid, a tissue extract, freshly harvested cells, or cells that have been incubated in tissue culture, in the presence of a detectably labeled antibody capable of identifying fingerprint gene peptides, and detecting the bound antibody by any of a number of techniques well known in the art.

The biological sample may be brought in contact with and immobilized onto a solid phase support or carrier such as nitrocellulose, or other solid support that is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the detectably labeled gene-specific antibody. The solid phase support may then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on solid support may then be detected by conventional means.

The terms "solid phase support or carrier" are intended to encompass any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention.

52

The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a

5   sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

The binding activity of a given lot of anti-wild type or -mutant fingerprint gene peptide antibody may be determined according to well known methods. Those skilled in

10   the art will be able to determine operative and optimal assay conditions for each determination by employing routine experimentation.

One of the ways in which the gene peptide-specific antibody can be detectably labeled is by linking the same to an enzyme and using it in an enzyme immunoassay (EIA) (Voller, *Ric Clin Lab*, 8:289-98 (1978) ["The Enzyme Linked Immunosorbent

15   Assay (ELISA)", Diagnostic Horizons 2:1-7, 1978, Microbiological Associates Quarterly Publication, Walkersville, Md.]; Voller, *et al.*, *J. Clin. Pathol.*, 31:507-20 (1978); Butler, *Meth. Enzymol.*, 73:482-523 (1981); Maggio (ed.), Enzyme Immunoassay, CRC Press, Boca Raton, Fla. (1980); Ishikawa, *et al.*, (eds.) Enzyme Immunoassay, Igaku-Shoin, Tokyo (1981)). The enzyme that is bound to the antibody will react with an appropriate

20   substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety that can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes that can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate

25   isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods that employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a

30   substrate in comparison with similarly prepared standards.

53

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect fingerprint gene wild type, mutant, or expanded peptides through the use of a radioimmunoassay (RIA) (*see, e.g.*, Weintraub, B.,

5   Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the

10   fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using fluorescence emitting metals

15   such as ^{152}Eu, or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediamine-tetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by

20   detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the antibody of the

25   present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

30   Throughout this application, various publications, patents and published patent applications are referred to by an identifying citation. The disclosures of these

54

publications, patents and published patent specifications referenced in this application are hereby incorporated by reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

The following examples are intended only to illustrate the present invention and
5   should in no way be construed as limiting the subject invention.

Examples

Example 1: Generation and Analysis of Mice Comprising TSH-R Gene Disruptions
10
To investigate the role of TSH-Rs, disruptions in TSH-R genes were produced by homologous recombination. Specifically, transgenic mice comprising disruptions in TSH-R genes were created. More particularly, as shown in Figure 2, a TSH-R-specific targeting construct having the ability to disrupt a TSH-R gene, specifically comprising
15   SEQ ID NO:1, was created using as the targeting arms (homologous sequences) in the construct, the oligonucleotide sequences identified herein as SEQ ID NO:3 or SEQ ID NO:4.

The targeting construct was introduced into ES cells derived from the 129/OlaHsd mouse substrain were used to generate chimeric mice. F1 mice were generated by
20   breeding with C57BL/6 females. The resultant F1N0 heterozygotes were backcrossed to C57BL/6 mice to generate F1N1 heterozygotes. F2N1 homozygous mutant mice were produced by intercrossing F1N1 heterozygous males and females.

The transgenic mice comprising disruptions in TSH-R genes were analyzed for phenotypic changes and expression patterns. The phenotypes associated with a
25   disruption in the TSH-R gene were determined. The homozygous mice demonstrated at least one of the following phenotypes:

Physical Examination

Mice were evaluated by physical examination and gross observation. At least one of the following phenotypes was observed in homozygous mutant mice:
30   ·   Body Shape: dwarfism
·   General: abnormal (small) and hunched posture
·   Head Shape/Snout: small head / short or abnormal snout

55

- Eyes and Ears: small
- Behavior/Locomotion: abnormal activity, posture, and/or responsiveness.

When compared to age- and gender-matched wild-type control mice, multiple parameters were abnormal in homozygous mutant mice. Five of six homozygous mice were dwarfs and the remaining homozygous female was classified as hunched, although its body weight and length were similar to those designated as dwarfs. All of the dwarfs had small heads, three had a short or abnormal snout and small eyes, and two had small ears. One female dwarf was non-responsive with abnormal posture and one male dwarf was hypoactive with abnormal posture. The homozygous female not identified as a dwarf was hyperactive.

Necropsy

Mice were necropsied, with body weight, body length, and organ weights obtained and gross pathological changes recorded. At least one of the following phenotypes was observed for homozygous mutant mice (see Figs 3 and 4):

- Thymus Gland: small
- Femur: malformed
- Skeletal Muscle: small
- Subcutis: decreased fat
- Seminal Vesicles: small or not visible
- Body Weight: low
- Body Length: short
- Organ Weight: low for spleen, liver, kidneys, heart, and thymus gland
- Organ Weight to Body Weight Ratio: low for spleen, liver, and kidneys

When compared to age- and gender-matched wild-type control mice, several gross abnormalities were noted in homozygous mutant mice. The thymus gland was small in two female and one male homozygote. Skeletal muscle was small in one female and one male homozygote, and this male also had decreased subcutaneous fat. The femur was malformed in two female and one male homozygote. The seminal vesicles were small or could not be located in all three homozygous males.

When compared to age- and gender-matched wild-type control mice, there were many differences in homozygous mutant mice that reflected their small size. Body

weights for homozygous females were about 51 to about 63 percent lower, and for homozygous males were about 47 to about 67 percent lower than wild-type mice. Body lengths for homozygous females were about 3 to about 30 percent lower, and for homozygous males were about 16 to about 28 percent lower than wild-type mice. Low organ weights were present in all homozygotes for spleen, liver, kidneys, and heart. Thymus gland weights were low in the four homozygotes for which thymic weights were obtained. Organ weight to body weight ratios for these organs tended to be low for the spleen, liver, and kidney in all or most homozygotes, but only in individual females for the thymus gland and heart.

Histopathology

Tissues from mice were evaluated histologically and at least one of the following phenotypes was observed in homozygous mutant mice:

- Thyroid Gland: small gland with small follicles
- Pituitary Gland: adenohypophysis, large and vacuolated cells, reduced chromophils; pars distalis, chromophobe hypertrophy
- Bone, Femur and Tibia (Stifle Joint): epiphyses, dysplasia / reduced and patchy ossification
- Bone Marrow: reduced cellularity
- Thymus Gland: hypoplasia with absence of cortico-medullary distinction
- Kidneys: immature with small glomeruli; lymphocytic infiltrates
- Testes: hypospermatogenesis / immature
- Testes: interstitial (Leydig) cell hypoplasia
- Epididymides: oligospermia
- Lung: lymphocytic infiltrates
- Eye: diffuse retinal fibrosis

When compared to age- and gender-matched wild-type control mice of about 49 days of age, most organs from homozygous mice were smaller in tissue sections, reflecting the smaller body size of these mice. Thyroid glands were very small with reduced follicle size in three of six homozygous mice. No thyroid gland was found in multiple sections taken from the other three homozygous mice, suggesting that they also had small thyroid glands. Multiple organs in homozygous mutant mice had additional

histologic alterations other than small size. The anterior pituitary gland (adenohypophysis) of all homozygous mice contained large vacuolated cells. The number of chromophils is diminished. Epiphyses of the femur and tibia (in the stifle joint section) had mild to severe dysplasia characterized by reduced and patchy ossification in

5   all homozygous mice. All homozygous mice had a minimal to moderate reduction in bone marrow cellularity involving both erythroid and myeloid cell series. There was moderate to severe thymic hypoplasia in three of six homozygous mice, without cortico-medullary distinction. Kidneys had small glomeruli reflecting minimal to moderate renal immaturity in all homozygotes. The testes appeared immature with hypospermatogenesis

10   in two of three homozygous males, and all homozygous males had mild to severe interstitial (Leydig) cell hypoplasia. The epididymides had moderate to severe oligospermia in all homozygous males, a sequel to testicular changes.

When compared to age- and gender-matched wild-type control mice of about 300 days of age, homozygous mutant mice had additional histologic alterations. The pituitary

15   glands exhibited hypertrophy of the chromophobes in three of four of these homozygous mutants. There was minimal lymphocytic infiltrate detected in the lungs of two of three of the about 300-day-old homozygous mutant mice. One female homozygous mutant mouse of about 300 days of age exhibited moderate, diffuse fibrosis of the retina. All homozygous mutant mice of about 300 days of age exhibited at least minimal

20   lymphocytic infiltrates in the kidneys.

Clinical Chemistry

Serum samples from mice were evaluated by a clinical chemistry panel and the homozygous mutant mice were found to exhibit: abnormal serum chemistry, including elevated blood urea nitrogen.

25   When compared to age- and gender-matched wild-type control mice, levels of blood urea nitrogen were elevated in the homozygous mutant mice (see Table 1 below).

TABLE 1

	Mouse	Age at Test (days)	Renal Function (Blood Urea Nitrogen) (mg/dL)
+/+	Female		
	83375	48	25.8
	114962	51	21.9
+/+	Male		
	83367	48	28.1
	83369	48	21.8
-/+	Female		
	83376	48	25.4
-/+	Male		
	83371	48	28.3
-/-	Female		
	90363	48	43.4
	113423	49	36.6
	113428	49	41.4
	119044	55	33
-/-	Male		
	102180	49	43.4
	102181	49	50.9
	113420	49	41.9
	113422	49	42.6
	113426	49	51.4
	113427	49	55.1
+/+	Female		
	107609	113	25.5
	107613	113	21.4
	107617	113	18.1
-/-	Female		
	111660	113	30.3
-/-	Male		
	107614	113	27.6
	111658	113	32.3
+/+	Female		
	107609	202	27.7
	107615	202	23.2
+/+	Male		
	107613	202	26.2
	107617	202	31.9
-/-	Female		

	111660	202	40.4
-/-	Male		
	107614	202	
	111658	202	32.3
	122137	203	37.3
	122138	203	31.7
+/+	Female		
	107609	308	28.7
	107615	308	21.1
+/+	Male		
	107613	308	17.1
	107617	308	25.1
-/-	Female		
	111660	307	38.8
-/-	Male		
	107614	307	35.6
	111658	307	33.5

Fertility

Homozygous mutant mice of each gender were set up in fertility matings at about

5   seven to eight weeks of age and found to exhibit a reproductive abnormality or infertility.
Specifically, homozygous mutant males and/or females were not fertile.

LacZ/RT-PCR Expression

RT-PCR Expression. Total RNA was isolated from the organs or tissues from
adult C57BL/6 wild type mice. RNA was DNaseI treated, and reverse transcribed using

10   random primers. The resulting cDNA was checked for the absence of genomic
contamination using primers specific to non-transcribed genomic mouse DNA. cDNAs
were balanced for concentration using HPRT primers. RNA transcripts were detectable
in brain, subcortical region, cerebellum, brainstem, spinal cord, eye, Harderian glands,
heart, lung, liver, pancreas, kidney, thymus, lymph nodes, bone marrow, skin,

15   gallbladder, urinary bladder, pituitary gland, adrenal gland, salivary gland, skeletal
muscle, tongue, stomach, small intestine, large intestine, testis, epididymis, seminal
vesicle, coagulating gland, prostate gland, ovaries and uterus. The highest level of RNA
transcripts was detected in white fat. No RNA transcripts were detectable in cortex,
olfactory bulb, spleen and cecum.

LacZ Reporter Gene Expression. In general, tissues from approximately 7-12 week old heterozygous mutant mice were analyzed for lacZ expression. Organs from heterozygous mutant mice were frozen, sectioned (about 10 μm), stained and analyzed for lacZ expression using X-Gal as a substrate for beta-galactosidase, followed by a

5   Nuclear Fast Red counterstaining.

In addition, for brain, wholemount staining was performed. The dissected brain was cut longitudinally, fixed and stained using X-Gal as the substrate for beta-galactosidase. The reaction was stopped by washing the brain in PBS and then fixed in PBS-buffered formaldehyde.

10   Wild-type control tissues were also stained for lacZ expression to reveal any background or signals due to endogenous beta-galactosidase activity. The following tissues can show staining in the wild-type control sections and are therefore not suitable for X-gal staining: small and large intestines, stomach, vas deferens and epididymis. It has been previously reported that these organs contain high levels of endogenous beta-

15   galactosidase activity.

LacZ (beta-galactosidase) expression was detectable in testis. LacZ expression was not detected in brain, spinal cord, sciatic nerve, eye, Harderian glands, thymus, spleen, lymph nodes, bone marrow, aorta, heart, lung, liver, gall bladder, pancreas, kidney, urinary bladder, trachea, larynx, esophagus, thyroid gland, pituitary gland,

20   adrenal glands, salivary glands, tongue, skeletal muscle, skin and female reproductive systems.

Expression:

Male Reproductive Systems

Testis

25   Many spermatogenic cells of the seminiferous tubules showed faint X-Gal staining.

In summary, when compared to age- and gender-matched wild-type control mice, multiple differences were present. At physical examination, the homozygotes were dwarfed/small, with small head, eyes, and ears, and abnormal snouts. Some exhibited

30   abnormal (hypo- or hyper-) activity, hunched posture, and/or non-responsiveness. Necropsy findings in the homozygotes included small thymus glands, small skeletal

61

muscle, reduced subcutaneous body fat, and seminal vesicles that were small or not located. Femurs were malformed. The homozygotes had low body weights and short body lengths. Low organ weights were present for spleen, liver, kidneys, heart, and thymus gland. Organ weight:body weight ratios were low for the spleen, liver, and

5   kidneys. Histopathological lesions were present in multiple organs in the homozygotes. Thyroid glands were small with small follicle size. The anterior pituitary gland had large and vacuolated cells, and decreased chromophils. The epiphyses of the femur and tibia had dysplasia characterized by reduced and patchy ossification. Bone marrow had reduced cellularity. The thymus gland was hypoplastic with absence of cortico-medullary

10   distinction. Kidneys were immature with small glomeruli. Testes were immature with hypospermatogenesis, and interstitial (Leydig) cells were hypoplastic. Epididymides had oligospermia. Blood urea nitrogen levels, were elevated. Homozygous mutant males and/or females were not fertile.

As is apparent to one of skill in the art, various modifications of the above

15   embodiments can be made without departing from the spirit and scope of this invention. These modifications and variations are within the scope of this invention.